arXiv:hep-ex/0103016v1  13 Mar 2001

# $B$ Mixing[*]

*Colin Gay*

Department of Physics, Yale University, New Haven, Connecticut 06511;
email: colin.gay@yale.edu

KEYWORDS:   $B$, oscillation, mixing, time-dependent

**ABSTRACT:**   The neutral $B$ mesons, $B^0$ and $B_s^0$, can oscillate between their particle and antiparticle states owing to flavor-changing weak interactions. In recent years, techniques to detect these oscillations as a function of the meson's decay time have been developed. In this article the physics of flavor oscillations is reviewed and theoretical predictions are summarized. The many observations that demonstrate the time-dependence of $B^0 - \overline{B^0}$ oscillations are presented along with a combined measurement of its frequency, $\Delta m_d = 0.484 \pm 0.015$ ps$^{-1}$. The attempts to measure the $B_s^0$ oscillation frequency, both directly and indirectly, are then summarized, currently resulting in a limit of $\Delta m_s > 14.6$ ps$^{-1}$ (95% CL). Finally, values for the CKM elements $|V_{td}| = (3.6 \pm 0.4) \times 10^{-3}$ and $|V_{ts}/V_{td}| > 4.7$ (95% CL) are extracted.

CONTENTS

---

[*]This paper is dedicated to the memory of Elizabeth Bishop Martin, who was originally intended to co-author this paper, but sadly passed away on March 16, 1999.

## 1   INTRODUCTION

The ability of a very few neutral mesons to change from their particle to their antiparticle state is a remarkable consequence of basic quantum mechanics and the structure of the weak interaction. This oscillation from matter to antimatter can be used to measure fundamental parameters of the standard model; in addition, it might have far reaching effects, such as breaking the matter/antimatter symmetry of the universe.

In 1955, Gell-Mann & Pais [1] showed that if a $\overline{K}^0$ meson with strangeness $S=-1$, as well as the $K^0$ meson with strangeness $S=1$ exists, then a quantum-mechanical mixing due to $K^0 - \overline{K}^0$ interactions takes place. It produces two physical particles, $K_1^0$ and $K_2^0$, that are a mixture of these states of well-defined strangeness, or flavor. Gell-Mann & Pais predicted the existence of a long-lived neutral kaon, $K_2^0$, which decays to three pions, as the companion to the shorter-lived $K_1^0$ that had already been observed. Lande [2] at Brookhaven confirmed the existence of the longer lived state, $K_L^0$, in 1956.

If the physical particles are a mixture of states of well-defined flavor, then these flavor eigenstates can be considered mixtures of the physical particles. These physical states must have slightly different masses, as discussed in the following section, and so they develop a phase difference as they evolve in time. Therefore the physical particle content of a flavor eigenstate evolves with time – an initially pure flavor eigenstate develops a component of the opposite flavor. The mixing of flavor eigenstates to form the physical particles, then, is equivalent to the oscillations of flavor eigenstates into one another.

The only hadrons that can undergo these oscillations are the following mesons: $K^0, D^0, B^0$ and $B_s^0$. The $\pi^0$ is its own antiparticle, the top quark is so heavy that it decays before forming stable hadrons, and excited meson states decay strongly or electromagnetically before any mixing can occur.

Such particle-antiparticle mixing has since been seen for $B$ mesons, first in an admixture of $B^0$ and $B_s^0$ by UA1 [3] and then in $B^0$ mesons by ARGUS [4] and later CLEO [5]. $B_s^0$ mixing was established by comparing the time-integrated oscillation probability for $B^0$ mesons, measured by ARGUS and CLEO, to that measured at LEP, which contains both $B^0$ and $B_s^0$ contributions [6]. Mixing is expected to be a very small effect in $D^0$ mesons and has not been observed. The experimental part of this review focuses specifically on recent analyses that attempt to measure the time-structure of $B^0$ and $B_s^0$ oscillations directly, rather than those that studied the oscillations in a time-integrated manner. The latter include those from ARGUS and CLEO [7] and earlier analyses from the experiments presented in this paper. See [6] for a list of these measurements.

The outline of this paper is as follows. Section 2 collects the various elements required to derive the mixing formalism; Section 2.3 presents results on quantities needed to extract Cabbibo-Kobayashi-Maskawa (CKM) matrix elements from mixing measurements; Section 3 presents an overview of experimental techniques used to measure the time dependence of the flavor oscillations; Section 4 presents

measurements of the $B^0$ oscillation frequency; Section 5 presents limits on direct searches for $B_s^0$ flavor oscillations; and Section 6 deals with measuring the $B_s^0$ oscillation frequency through lifetime differences in the $B_s^0$ system.

Since this paper's focus is $B$-meson mixing, the formalism is presented in terms of $b$ quarks coupled with $d, s$ quarks, though it was first derived for the kaon system. $B$ is used generically to mean either $B^0$ or $B_s^0$ in cases where the result applies to both; however for clarity and brevity only the $B^0$ Feynman diagrams are shown. Also, the charge conjugate of listed decay modes is implied unless explicitly stated otherwise. For convenience we have set $\hbar = c = 1$ and supress the $c$ in the units of momentum and mass.

## 2 MIXING FORMALISM

Particle-antiparticle oscillations are possible because of the flavor-changing term of the standard model Lagrangian,

$$\mathcal{L} = \frac{g}{\sqrt{2}} (\overline{u}, \overline{c}, \overline{t})_L V_{\text{CKM}} \gamma_\mu \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L W^\mu + h.c.,$$

where $V_{\text{CKM}}$ is the CKM matrix [8]. A popular parameterization of this matrix is that of Wolfenstein [9], which expands each term in powers of the Cabibbo angle $\lambda \approx 0.22$, shown below to $\mathcal{O}(\lambda^4)$:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda(1 + iA^2\lambda^4\eta) & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}.$$

This term in the Lagrangian engenders box diagrams involving an internal loop with two $W$ bosons, of the sort shown in Figure 2, which result in nonzero transition matrix elements between $B^0 - \overline{B}^0$ and $B_s^0 - \overline{B_s^0}$.

Before calculating these transition amplitudes within the standard model, the relevant general formalism is developed, i.e. the quantum mechanics of a two-state, particle-antiparticle system weakly coupled to a continuum.

### 2.1 Introduction to Mixing Formalism

The oscillation effect follows from a simple perturbative solution to Schrödinger's equation. Let $\mathcal{H}_0$ be the Hamiltonian of the strong interaction. If this were the only force, there would be stable states, $|B >= |\overline{b}q >$ and $|\overline{B} >= |b\overline{q} >$ $(q = d, s)$, that are eigenvectors of $\mathcal{H}_0$. By the CPT theorem, the masses are equal, $m_B = m_{\overline{B}} = m_0$, and the Hamiltonian is

$$\mathcal{H}_0 = \begin{pmatrix} m_0 & 0 \\ 0 & m_0 \end{pmatrix}.$$

When the weak interaction $\mathcal{H}_W$ is added, the simple two-state system becomes much more complicated, as shown in Figure 1. The weak force is responsible for nonzero matrix elements between the two states to a continuum of states (i.e. possible decay modes), which can be split into three groups: $(a)$ states accessible only to $|B>$, i.e. states $|\alpha >$ with $<\alpha|\mathcal{H}_W|\overline{B}>= 0$; $(b)$ states accessible only to $|\overline{B}>$; and $(c)$ states coupled to both $|B>$ and $|\overline{B}>$. In addition, the two discrete

states are connected by a direct matrix element $W_{12} = <\overline{B}|\mathcal{H}_W|B>$ or via off-shell (i.e. with $E \neq m_0$) continuum states accessible to both. The Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_W$ has infinite dimensions.

**\<B|W|B\>**

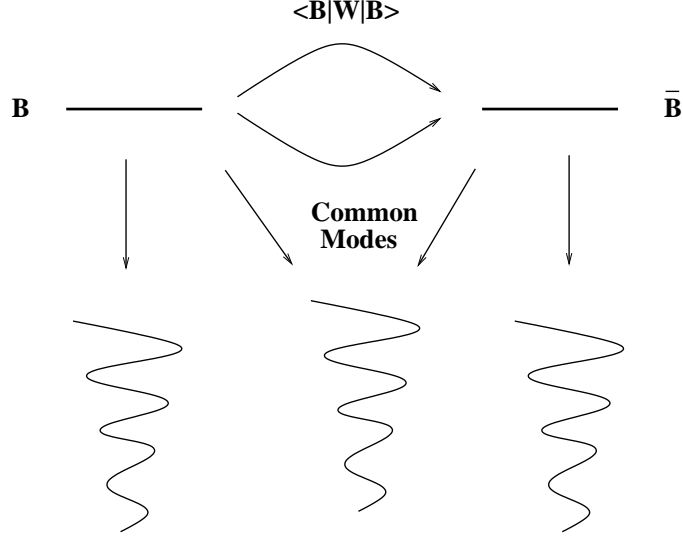**B** ——————                ——————  **B̄**

**Common
Modes**

Figure 1: Schematic of a two-state system coupled weakly to a continuum (indicated by the wiggles)

The time evolution of a wavefunction that is a pure $|B>$ at $t = 0$ is easily calculated using standard first-order perturbation theory, and in general the wavefunction develops both $|\overline{B}>$ and continuum components. Typically, one limits the state space to that spanned by the $|B>$ and $|\overline{B}>$, however, and derives a matrix $\mathcal{H}$ (still referred to as a Hamiltonian even though it is clearly not hermitian) which allows one to use the Schrödinger equation in this subspace. It is easy to show it has the form:

$$\mathcal{H} = \begin{pmatrix} m_0 + \delta E & W_{12} + \delta E_{12} \\ W_{12}^* + \delta E_{12}^* & m_0 + \delta E \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix}. \tag{1}$$

The introduction of the continuum states has several effects:

1. It introduces a slight mass shift $\delta E$ in both $|B>$ and $|\overline{B}>$, which can be absorbed into the mass $M = m_0 + \delta E$.

2. It introduces off-diagonal elements $M_{12} = W_{12} + \delta E_{12}$ in the real part of $\mathcal{H}$ due to the coupling of the particle and antiparticle via $\mathcal{H}_W$ and via *off-shell* continuum states accessible to both $|B>$ and $|\overline{B}> (\delta E_{12})$.

3. It introduces an imaginary part to $\mathcal{H}$, with diagonal elements given by the matrix elements of $|B>$ and $|\overline{B}>$ to the on-shell continuum states unique to each.

4. It introduces off-diagonal elements $\Gamma_{12}$ into the imaginary part of $\mathcal{H}$ given by the matrix elements of $|B>$ and $|\overline{B}>$ to the on-shell continuum states common to both.

Diagonalizing this $\mathcal{H}$ gives us the eigenvectors and eigenvalues. The eigenstates are conventionally denoted $H, L$, for heavy and light, rather than short and long as customary for the kaon system. If

$$
\begin{aligned}
|B_L> &= p|B> +q|\overline{B}> \qquad \text{and} \\
|B_H> &= p|B> -q|\overline{B}>,
\end{aligned}
\tag{2}
$$

then

$$
\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}},
\tag{3}
$$

and the eigenvalues are

$$
\lambda_{H,L} = m_{H,L} - \frac{i}{2}\Gamma_{H,L},
\tag{4}
$$

where the masses and widths of these states are

$$
m_H, m_L = M \pm Re\sqrt{|M_{12}|^2 - \frac{|\Gamma_{12}|^2}{4} - iRe(M_{12}\Gamma_{12}^*)} \equiv M \pm \Delta m/2,
$$

$$
\Gamma_H, \Gamma_L = \Gamma \pm 2Im\sqrt{|M_{12}|^2 - \frac{|\Gamma_{12}|^2}{4} - iRe(M_{12}\Gamma_{12}^*)} \equiv \Gamma \pm \Delta\Gamma/2,
\tag{5}
$$

which satisfy

$$
\begin{aligned}
\Delta m^2 - \frac{\Delta\Gamma^2}{4} &= 4(|M_{12}|^2 - \frac{|\Gamma_{12}|^2}{4}) \qquad \text{and} \\
\Delta m \Delta\Gamma &= 4Re(M_{12}\Gamma_{12}^*).
\end{aligned}
\tag{6}
$$

Equation 2 can be rearranged to give the flavor eigenstate $|B>$ in terms of $|B_H>, |B_L>$. The standard derivation uses their simple time dependence, $[|B_H(t)> = e^{-i\lambda_H t}|B_H>]$, to obtain the time evolution of a state which was a pure $|B>$ at $t = 0$, expressed in terms of the flavor eigenstates. (This procedure implies the existence of a reference frame with both the $|B_L>$ and $|B_H>$ components of the $|B>$ at rest, and such a frame does not, strictly speaking, exist. However, setting the relative velocity of the $|B_L>$ and $|B_H>$ states to zero is an excellent approximation. A comment on its validity and the appropriateness of describing the system's evolution as a function of time can be found below.)

$$
\begin{aligned}
|B(t)> &= \frac{1}{2p}(|B_L(t)> +|B_H(t)>) \\
&= \frac{1}{2}e^{-iMt}e^{-\frac{\Gamma}{2}t}\left((e^{\frac{\Delta\Gamma}{4}t}e^{i\frac{\Delta m}{2}t} + e^{-\frac{\Delta\Gamma}{4}t}e^{-i\frac{\Delta m}{2}t})|B> \right. \\
&\qquad \left. + \frac{q}{p}(e^{\frac{\Delta\Gamma}{4}t}e^{i\frac{\Delta m}{2}t} - e^{-\frac{\Delta\Gamma}{4}t}e^{-i\frac{\Delta m}{2}t})|\overline{B}>\right).
\end{aligned}
\tag{7}
$$

Similarly, the time evolution of a state which is pure $|\overline{B}>$ as $t = 0$ is

$$
\begin{aligned}
|\overline{B}(t)> &= \frac{1}{2q}\left[|B_L(t)> -|B_H(t)>\right] \\
&= \frac{1}{2}e^{-iMt}e^{-\frac{\Gamma}{2}t}\left[\frac{p}{q}(e^{\frac{\Delta\Gamma}{4}t}e^{i\frac{\Delta m}{2}t} - e^{-\frac{\Delta\Gamma}{4}t}e^{-i\frac{\Delta m}{2}t})|B> \right. \\
&\qquad \left. + (e^{\frac{\Delta\Gamma}{4}t}e^{i\frac{\Delta m}{2}t} + e^{-\frac{\Delta\Gamma}{4}t}e^{-i\frac{\Delta m}{2}t})|\overline{B}>\right].
\end{aligned}
\tag{8}
$$

The normalizations of these states are given by

$$\eta^2 = \int_0^\infty <B(t)|B(t)> dt = \frac{\Gamma}{2}\left[\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4}+\frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right]$$

$$\overline{\eta}^2 = \int_0^\infty <\overline{B}(t)|\overline{B}(t)> dt = \frac{\Gamma}{2}|\frac{p}{q}|^2\left[\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4}-\frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right]. \qquad (9)$$

Let $P_m^B(t)$ denote the probability that a particle produced as a $B$ oscillated (mixed) and decayed as a $\overline{B}$. [i.e. $P_m^B(t) = \frac{1}{\eta^2}| <\overline{B}|B(t)> |^2$.] Let $P_u^B(t)$ denote the conjugate probability that this particle did not oscillate, that is, it remained unmixed (with similar definitions for initial $\overline{B}$ states). Then Equations 7, 8 and 9 give the following:

$$P_u^B(t) = \frac{e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4}+\frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t + \cos\Delta mt),$$

$$P_m^B(t) = \frac{|q/p|^2 e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4}+\frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t - \cos\Delta mt), \qquad (10)$$

$$P_u^{\overline{B}}(t) = \frac{|q/p|^2 e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4}-\frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t + \cos\Delta mt),$$

$$P_m^{\overline{B}}(t) = \frac{e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4}-\frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t - \cos\Delta mt). \qquad (11)$$

Note that these expressions are not symmetric between $B$ and $\overline{B}$ states.

These formulae have two limiting cases: neglecting CP violation in the mixing, and neglecting the lifetime difference $\Delta\Gamma$ (which also in general implies there is no CP violation in the mixing).

Equation 2 can be rewritten as (similarly for the $|B_H>$)

$$|B_L> = \frac{p+q}{2}\left[(|B> +|\overline{B}>) + \frac{1-q/p}{1+q/p}(|B> -|\overline{B}>)\right],$$

and thus $(1-q/p)/(1+q/p) \equiv \epsilon_B$ is a measure of the amount by which $|B_H>$ and $|B_L>$ differ from CP eigenstates. $\epsilon_B$ is expected to be very small in the standard model, $\mathcal{O}(10^{-3})$. The current world average for the $B^0$ system is $\text{Re}(\epsilon_B)=0.002\pm 0.007$ [6]. No measurement or limit exists for the $B_s^0$ system. The limit of no CP violation in mixing is thus $q/p = 1$. In this limit the $B$ and $\overline{B}$ symmetry is regained, and we obtain unmixed and mixed decay probabilities for both $B$ and $\overline{B}$ of:

$$P_{u,m}(t) = \frac{1}{2}\Gamma e^{-\Gamma t}\left(1-\frac{\Delta\Gamma^2}{4\Gamma^2}\right)(\cosh\frac{\Delta\Gamma}{2}t \pm \cos\Delta mt), \qquad (12)$$

where the $+$ sign corresponds to $P_u$. This form is appropriate for $B_s^0$ mesons, which are not expected to be subject to large CP-violating effects.

On the other hand, even in the presence of CP violation, a simple form can be obtained. The lifetime difference between the heavy and light states is expected to be small, $\Delta\Gamma/\Gamma \leq 1\%$ for the $B^0$ and perhaps as large as 25% for the $B_s^0$ [10].

From Equation 6, $\Delta\Gamma = 0$ in general only if $\Gamma_{12} = 0$. In this case,

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} = \sqrt{\frac{M_{12}^*}{M_{12}}} = e^{-i\phi},$$

thus $|q/p| = 1$. In this $\Delta\Gamma = 0$ limit, the time evolutions from Equations 7 and 8 become

$$|B(t) > = e^{-iMt}e^{-\frac{\Gamma}{2}t}\left(\cos\frac{\Delta m}{2}t|B> +ie^{-i\phi}\sin\frac{\Delta m}{2}t|\overline{B}>\right)$$

$$|\overline{B}(t) > = e^{-iMt}e^{-\frac{\Gamma}{2}t}\left(\cos\frac{\Delta m}{2}t|\overline{B}> +ie^{+i\phi}\sin\frac{\Delta m}{2}t|B>\right). \qquad (13)$$

The mixed and unmixed decay probabilities again become equal for the $B$ and $\overline{B}$ mesons:

$$P_{u,m}(t) = \frac{1}{2}\Gamma e^{-\Gamma t}(1 \pm \cos\Delta mt). \qquad (14)$$

This form is expected to be appropriate for $B^0$ mesons, for which a large phase $\phi$ (the source of mixing-induced CP violation) is possible.

Equations 12 and 14 will be the expressions used throughout this paper. Their time-integrated versions, expressing the probability that a $B$ decays as a $\overline{B}$, are (with $x \equiv \Delta m/\Gamma$)

$$\chi = \int_0^\infty P_m(t) = \frac{1}{2}\frac{x^2 + \frac{1}{4}\frac{\Delta\Gamma^2}{\Gamma^2}}{1 + x^2} \qquad (15)$$

and in the $\Delta\Gamma = 0$ limit,

$$\chi = \frac{1}{2}\frac{x^2}{1 + x^2}. \qquad (16)$$

As has been pointed out in the context of neutrino oscillation experiments [11], the oscillations observed by any experiment are oscillations in space not in time. That is, one has a source creating a pure $B$ or $\overline{B}$ meson, which may have oscillated by the time it reaches a distant detector. In this spatial picture, we have a source, very small compared to the oscillation wavelength, which emits a pure $B$ meson. The boundary condition that must be imposed, then, is that the probability of finding a $\overline{B}$ meson at the source must vanish for all time, otherwise a pure $B$ would not be emanating. The $|B_L>$ and $|B_H>$ components propagate with phase $e^{i(E_{L,H}t-p_{L,H}x)}$, where $x$ denotes the direction of motion. At the origin, the only way to ensure the wavefuntion does not change the relative $|B_L> -|B_H>$ phase and develop a $\overline{B}$ component is the condition $E_L = E_H$. That is, the $B$ meson has a definite energy. The components $|B_L>$ and $|B_H>$ will have the same energy, but different momenta $p_{L,H} = \sqrt{E^2 - m_{L,H}^2}$ respectively. This induces spatial oscillations that go as $e^{i(p_H-p_L)x}$. Thus, if the light and heavy components have the same energy but different momenta, they must be moving with different velocities, and there is no frame in which both the $|B_L>$ and $|B_H>$ are at rest. In the derivation above, however, we put $E_{L,H} = m_{L,H}$ for both states.

At the time $t$, the centers of the $|B_L>$ and $|B_H>$ wave packets have separated by a distance $d = (v_L - v_H)t = (p_L - p_H)t/E$. As $\Delta m$ is very small compared to the mean mass $M$ ($3 \times 10^{-13}$ GeV vs. 5.3 GeV), this distance is $d \approx \frac{2M\Delta m}{EP}ct$, where $P^2 = E^2 - M^2$. Setting $t$ as large as 10 ps (the $B$ lifetime is approximately

1.5 ps), and noting that the typical $B$-meson energy in the current experiments is $10 - 30$ GeV, we see that the wavepacket separation is only $\mathcal{O}(10^{-10}\mu\text{m})$. This justifies treating the system as moving with a single group velocity given by the mean velocity $v = P/E$. Thus the spatial dependence of the oscillations, $(p_L - p_H)x$, can be converted to a time dependence using $x = vt = P/E \approx (p_L+p_H)/2E$ to obtain $(p_L^2 - p_H^2)t/2E = (m_L^2 - m_H^2)t/2E$. If we substitute $E = \gamma M$, and the proper time $\tau = t/\gamma$, the time dependence goes as $(m_L^2 - m_H^2)\tau/2M = \Delta m\tau$, recovering the form used in the standard treatment above.

## 2.2  Mixing in the Standard Model

The time-evolution and decay probabilities presented in the previous section are valid for any perturbation to the strong Hamiltonian. This section evaluates the matrix elements $M_{12}$ and $\Gamma_{12}$, in the context of the standard model.
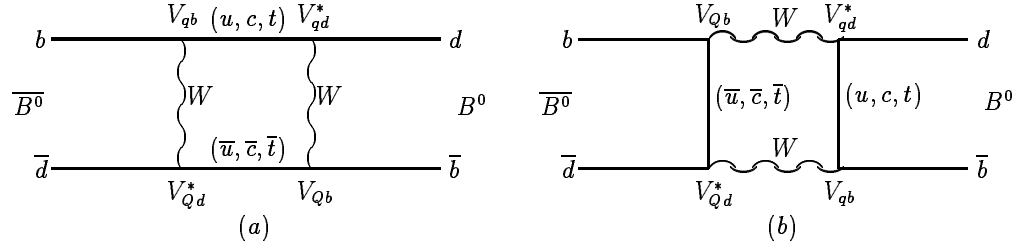


Figure 2: Lowest order box diagrams responsible for $B^0 - \overline{B^0}$ oscillations.

To lowest order, the matrix element coupling a $B^0$ and $\overline{B^0}$ is given by the Feynman diagrams shown in Figure 2$a,b$, with similar diagrams for the $B_s^0$. With three generations, the internal quark lines $Q, q$ can be $u, c$, or $t$. Of course we must take into account that the quark lines coming in and going out are not free – the $b$ and light quark $d$ or $s$ are bound into a hadron. These diagrams are evaluated using a method first employed by Gaillard & Lee [12] in 1974 for the $K^0$ system. The Hamiltonian corresponding to the box diagrams is sandwiched between the bound-state hadrons, $\mathcal{M} =< B|H_W|\overline{B} >$, (so that $M_{12} = \mathcal{M}/(2m_B)$, including the normalization factor), the momenta and masses of the incoming and outgoing quarks are set to zero (since they are small compared with $M_W$), and the internal loop momentum integral is performed.

Some care must be taken in evaluating these diagrams. A good choice of gauge to perform the calculation is the generalized renormalizable gauge, $R_\xi$, with $\xi$ finite, introduced by Fujikawa et al [13]. In the standard model, there is an unphysical charged scalar $\phi^\pm$, which is a remnant of the charged Higgs that is absorbed to give the $W$ boson its mass. In fact, such particles will generally exist in any model which generates masses for the vector bosons through spontaneous symmetry breaking. These scalars couple to quarks with a strength proportional to the masses. For an incoming $d$-type quark of generation $j$ and outgoing $u$-type of generation $i$, the vertex factor is $\frac{-ig}{2\sqrt{2}M_W}[m_{d_j}(1 + \gamma_5) - m_{u_j}(1 - \gamma_5)]V_{u_id_j}$.
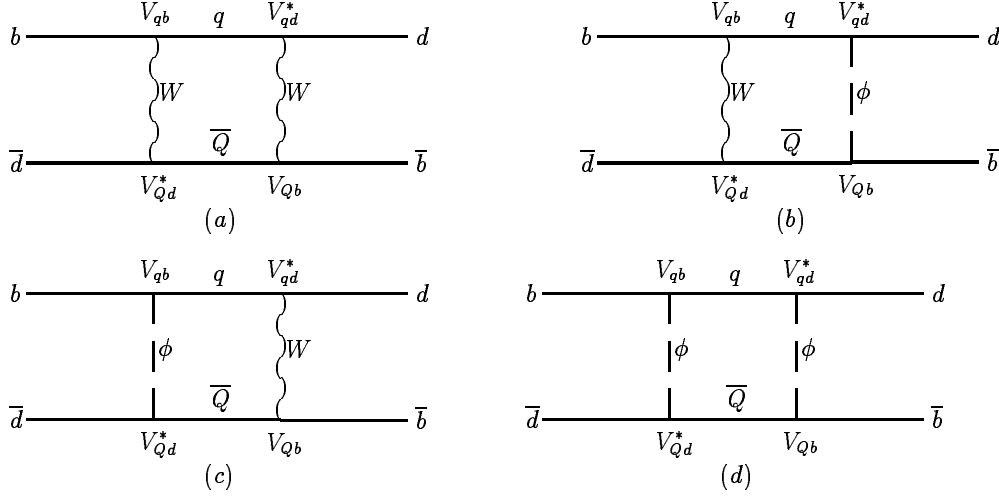
Figure 3: Examples of box diagrams for $b - \bar{b}$ oscillations, including unphysical scalar contributions.

The $W$ and $\phi$ propagators in the $R_\xi$ gauge are

$$W^\pm : \qquad \frac{-i}{k^2 - M_W^2 + i\epsilon} \left[ g_{\mu\nu} + \frac{(\xi - 1)k_\mu k_\nu}{k^2 - \xi M_W^2} \right],$$

$$\phi^\pm : \qquad \frac{i}{k^2 - \xi M_W^2 + i\epsilon}.$$

While the unitary gauge, $R_\infty$, is convenient for evaluating tree-level diagrams (the unphysical scalars drop out), in this gauge there are extraneous singularities generated in the box diagram from the $W$-propagator terms, which go as $k_\mu k_\nu / M_W^2$. In fact, the Green's functions are unrenormalizable, though one can show that the full S-matrix is finite [13], so extreme care must be taken in performing the $\xi \to \infty$ limit.

In the discussion below we work in the 't Hooft-Feynman gauge, $\xi = 1$, in which the propagators are particularly simple. Note that the relative minus sign in the propagators causes the diagrams in Figures 3b, c to add to, rather than cancel, the diagrams in Figures 3a, d.

For example, the basic matrix element for $B^0 - \overline{B^0}$ transitions through the diagram in Figure 3a, involving internal quarks of type $Q$ and $q$ is

$$
\begin{aligned}
i\mathcal{M} \;=\; & <B^0| \frac{g^4}{64}\xi_Q\xi_q \int \frac{d^4k}{(2\pi)^4} \overline{d}(p_3)\gamma^\mu(1 - \gamma_5)\frac{1}{\slashed{k} - m_Q}\gamma^\nu(1 - \gamma_5)b(p_1) \\
\times \; & \overline{d}(p_2)\gamma_\nu(1 - \gamma_5)\frac{1}{\slashed{k} - \slashed{p}_1 - \slashed{p}_2 - m_q}\gamma_\mu(1 - \gamma_5)b(p_4) \\
\times \; & \frac{1}{(k - p_1)^2 - M_W^2}\frac{1}{(k - p_3)^2 - M_W^2}|\overline{B^0}> .
\end{aligned}
$$

Here $(p_1, p_2)$ are the incoming $(b, \overline{d})$ momenta; $(p_3, p_4)$ are the outgoing $(d, \overline{b})$ momenta, and $\xi_Q = V_{Qd}^* V_{Qb}$ and $\xi_q = V_{qd}^* V_{qb}$ are the CKM factors.

We now invoke the approximation that the external $b$ and $d$ quark momenta and masses are zero, $p_i = 0$. After a reduction of the Dirac matrices, the matrix

element becomes

$$
i\mathcal{M}_{WW} = \frac{g^4}{16}\xi_Q\xi_q[< B^0|\overline{d}\gamma^\mu(1-\gamma_5)b\overline{d}\gamma_\mu(1-\gamma_5)b|\overline{B^0} >]
$$
$$
\times \int \frac{d^4k}{(2\pi)^4}\frac{k^2}{(k^2-m_Q^2)(k^2-m_q^2)(k^2-M_W^2)^2}.
$$

Applying the same assumptions to the diagrams with the scalars, one finds that all diagrams reduce to having a common operator structure, $\hat{O}_{LL} = (\overline{d}_L\gamma^\mu b_L)^2$, multiplied by a convergent integral over the loop momentum. The result is

$$
\mathcal{M}_{AB}(x_q, x_Q) = \frac{g^4}{64\pi^2 M_W^2}E_{AB}(x_q, x_Q) < B^0|(\overline{d}_L\gamma^\mu b_L)^2|\overline{B^0} >
$$

with $x_q \equiv m_q^2/m_W^2$ and

$$
E_{WW}(x_q, x_Q) = -\frac{1}{2}\frac{1}{x_q - x_Q}\left[\frac{x_q^2\log x_q}{(1-x_q)^2} - \frac{x_Q^2\log x_Q}{(1-x_Q)^2} + \frac{1}{1-x_q} - \frac{1}{1-x_Q}\right]
$$

$$
E_{W\phi}(x_q, x_Q) = \frac{1}{2}\frac{x_q x_Q}{x_q - x_Q}\left[\frac{x_q\log x_q}{(1-x_q)^2} - \frac{x_Q\log x_Q}{(1-x_Q)^2} + \frac{1}{1-x_q} - \frac{1}{1-x_Q}\right]
$$

$$
= E_{\phi W}(x_q, x_Q)
$$

$$
E_{\phi\phi}(x_q, x_Q) = \frac{x_q x_Q}{4}E_{WW}(x_q, x_Q).
$$

Table 1 provides values of these functions for $m_u = 0, m_c = 1.5$ GeV, and $m_t = 170$ GeV.

Table 1: Contributions of the $W$-$W$, $W$-scalar and scalar-scalar diagrams with all combinations of internal quark lines.

| Diagram | | **u, u** | **u, c** | **u, t** | **c, c** | **c, t** | **t, t** |
|---|---|---|---|---|---|---|---|
| (a) | $W-W$ | 1 | 1.9952 | 0.5353 | 0.9955 | 0.5352 | 0.1339 |
| (b)+(c) | $W-\phi$ | 0 | 0 | 0 | $1.45 \times 10^{-6}$ | $1.02 \times 10^{-3}$ | 1.196 |
| (d) | $\phi-\phi$ | 0 | 0 | 0 | $3.02 \times 10^{-8}$ | $2.08 \times 10^{-4}$ | 0.6690 |

Because the scalar couples to the quarks with a strength proportional to $m_q/M_W$, only the diagrams with two top quarks contribute substantially to the diagrams in Figures 3$b,c,d$. Although the $W - W$ diagrams with $u$ quarks and $c$ quarks have large matrix elements, they very nearly cancel when combined with their CKM factors because both the $u$ and $c$ quark are light compared to the $W$. Traditionally, the unitarity of the CKM matrix is assumed, and is used to express the $u$-quark terms $V_{ud}V_{ub}^* = -V_{cd}V_{cb}^* - V_{td}V_{tb}^*$. For example, it is not the case that the $ut$ matrix element is small, but rather its contribution almost exactly cancels the $ct$ contribution via this GIM-like mechanism [14] if the CKM matrix is unitary.

Thus the full matrix element, summing over all internal quark types, becomes simply a sum over $c$ and $t$ quarks and different bosons:

$$
\mathcal{M} = < B^0|(\overline{d}_L\gamma^\mu b_L)^2|\overline{B^0} > \frac{g^4}{64\pi^2 M_W^2}2\sum_{q,Q=c,t}\xi_Q\xi_q \times
$$

$$\sum_{A,B=W,\phi} \left[ E_{AB}(x_q, x_Q) - E_{AB}(x_q, 0) - E_{AB}(0, x_Q) + E_{AB}(0,0) \right].$$

To summarize, we have made the following assumptions: $(a)$ The $W$ mass is generated by a spontaneous symmetry-breaking mechanism, resulting in scalar remnants; $(b)$ the momenta and masses of the $b$ and $d$ quarks inside a bound meson can be neglected; and $(c)$ the CKM matrix is unitary.

This program was first carried out, with the important scalar terms included, by Inami & Lim [15] in 1981. The result of the sum over bosons is

$$f_3(x_Q, x_q) = \frac{x_Q x_q}{x_Q - x_q} \left( \ln \frac{x_Q}{x_q} - \frac{3}{4} \frac{x_Q^2 \ln x_Q}{(1 - x_Q)^2} + \frac{3}{4} \frac{x_q^2 \ln x_q}{(1 - x_q)^2} \right) - \frac{3}{4} \frac{x_Q x_q}{(1 - x_Q)(1 - x_q)}. \tag{17}$$

For the same quark type on the two internal lines $(q = Q)$, this reduces to

$$f_2(x_q) = x_q \left( \frac{1}{4} + \frac{9}{4} \frac{1}{1 - x_q} - \frac{3}{2} \frac{1}{(1 - x_q)^2} \right) - \frac{3}{2} \frac{x_q^3 \log x_q}{(1 - x_q)^3}. \tag{18}$$

Table 2 shows the size of the Inami-Lim functions $f_2(x_c)$, $f_2(x_t)$ and $f_3(x_t, x_c)$ $(= f_3(x_c, x_t))$, and the CKM factors that multiply each term. For the cases of $B^0$ and $B_s^0$ mesons, it turns out that these CKM factors are almost constant, independent of the type of quarks in the internal loop. For reference, the CKM coefficients for the $K^0$ system are shown, which vary by five orders of magnitude.

Table 2: Factors entering the matrix element, which is proportional to the product of the Inami-Lim function and the CKM term.

| Internal quarks | I-L factor | $B^0$ CKM | $B_s^0$ CKM | $K^0$ CKM |
|---|---|---|---|---|
| **c,c** | $3.5 \times 10^{-4}$ | $A^2\lambda^6$ $(7.4 \times 10^{-5})$ | $A^2\lambda^4$ $(1.4 \times 10^{-3})$ | $\lambda^2$ $(2.7 \times 10^{-2})$ |
| **c,t** | $3.0 \times 10^{-3}$ | $A^2\lambda^6|1 - \rho - i\eta|$ $(7.3 \times 10^{-5})$ | $A^2\lambda^4$ $(1.5 \times 10^{-3})$ | $A^2\lambda^6|1 - \rho - i\eta|$ $(8.8 \times 10^{-6})$ |
| **t,t** | $2.5$ | $A^2\lambda^6|1 - \rho - i\eta|^2$ $(7.2 \times 10^{-5})$ | $A^2\lambda^4$ $(1.5 \times 10^{-3})$ | $A^4\lambda^{10}|1 - \rho - i\eta|^2$ $(1.1 \times 10^{-7})$ |

Note that the function $f_2(x_q)$, for which both internal quarks are of the same type, is proportional to $x_q$, i.e. the square of the mass of the internal quark, and so the top quark loop completely dominates the charm loop's contribution. The function $f_3(x_t, x_c)$ is proportional to $x_c$ and is relatively small, an $\approx 0.1\%$ correction.

In the approximation that only the top quark contributes, the matrix element can thus be written (with a factor for all permutations) as

$$\mathcal{M} = \frac{G_F^2}{8\pi^2} |V_{td}^* V_{tb}|^2 m_t^2 f_2(m_t^2/M_W^2) <B^0|(\bar{d}_L \gamma^\mu b_L)^2|\overline{B^0}> . \tag{19}$$

Hence, we need to evaluate the effective matrix element

$$< B|(\overline{d}\gamma^\mu \quad (1 - \gamma_5) \quad b)^2|\overline{B}> =$$
$$\sum_n < B|\overline{d}\gamma^\mu(1 - \gamma_5)b|n >< n|\overline{d}\gamma_\mu(1 - \gamma_5)b|\overline{B} >$$
$$\equiv B_B < B|\overline{d}\gamma^\mu(1 - \gamma_5)b|0 >< 0|\overline{d}\gamma_\mu(1 - \gamma_5)b|\overline{B} >$$
$$= B_B| < 0|\overline{d}\gamma^\mu\gamma_5 b|B > |^2, \tag{20}$$

where $B_B$ is the bag factor, which is 1 if the vacuum insertion $|0 >< 0|$ saturates the sum over all intermediate states.

Substituting $< 0|\overline{d}\gamma^\mu\gamma_5 b|B >= if_B p^\mu$, where $f_B$ is the $B$ decay constant into this expression, and taking into account the three colours of the quarks gives

$$< B|(\overline{d}\gamma^\mu(1 - \gamma_5)b)^2|\overline{B} > \quad = \quad \frac{8}{3}B_B f_B^2 m_B^2. \tag{21}$$

Finally, the diagrams in Figure 3 can have arbitrary numbers of gluons running between any of the quark lines. The effect of this on $\mathcal{M}$ is encoded in the QCD correction factor $\eta_B$, which multiplies the $\mathcal{M}$ above.

The size of this correction, including next-to-leading-order (NLO) diagrams, has been estimated by Buras et al [16],

$$\eta_B \quad = \quad 0.55 \pm 0.01. \tag{22}$$

These authors found that $\eta_B$ depends upon the definition of the top quark mass and that only the product $\eta_B f_2(x_t)$ is insensitive to this choice. To be consistent with this estimate of $\eta_B$, the top quark mass must be chosen to be $\overline{m_t}(m_t^{\text{pole}})$, the mass calculated using the $\overline{MS}$ renormalization scheme evaluated at the measured (pole) mass of 174±5 GeV [6]. This gives a mass 6-7 GeV lower than the measured mass. Some confusion over $\eta_B$ exists in the literature; however $\eta_B$, by definition, is the same for $B^0$ and $B_s^0$ mesons.

Putting together Equations 19, 21, 22 results in

$$|M_{12}| = \frac{|\mathcal{M}|}{2m_B} \quad = \quad \frac{G_F^2}{6\pi^2}m_B f_B^2 B_B \eta_B m_t^2 f_2(m_t^2/M_W^2)|V_{td}^* V_{tb}|^2.$$

To relate this to the heavy-light mass difference $\Delta m$, we in principle also need the absorptive part of the matrix element, $\Gamma_{12}$, as per Equation 5. However, $\Gamma_{12} \ll M_{12}$ as argued below, and can be neglected.

The matrix element $\Gamma_{12}$ involves decays to modes common to both $|B >$ and $|\overline{B} >$ that are on-shell and thus energetically allowed, meaning the top quark loops do not contribute. Since the matrix element is given approximately by the square of the mass of the state in the loop, and this is $\approx m_b^2$ for on-shell transitions, we obtain $\Gamma_{12}/M_{12} \approx m_b^2/m_t^2 \ll 1$.

Alternatively, this can be seen dirctly from the data. The fractional lifetime difference $\Delta\Gamma/\Gamma$ is expected to be small, as previously mentioned – less than 0.01 for the $B^0$ and up to 0.25 for the $B_s^0$. As shown in Sections 4 and 5, current experiments demonstrate that $\Delta m \approx 0.7\Gamma$ for the $B^0$ system, and $\Delta m_s \geq 20\Gamma$ for the $B_s^0$. Combining this information gives $\Delta\Gamma_d = (0.01/0.7)\Delta m_d$ and $\Delta\Gamma_s = (0.25/20)\Delta m_s$, so in both cases $\Delta\Gamma \sim \mathcal{O}(10^{-2})\Delta m$. Equation 6 then implies, in general, that $|\Gamma_{12}| = \mathcal{O}(10^{-2})|M_{12}|$, leaving

$$\Delta m \quad = \quad 2|M_{12}| \qquad \text{and}$$
$$\Delta\Gamma \quad = \quad 2Re(M_{12}\Gamma_{12}^*)/|M_{12}|. \tag{23}$$

Hence, the final formulae are as follows:

$$
\begin{aligned}
\Delta m_d &= \frac{G_F^2}{6\pi^2} m_{B_d} f_{B_d}^2 B_{B_d} \eta_B m_t^2 f_2(m_t^2/M_W^2)|V_{td}^* V_{tb}|^2, \\
\Delta m_s &= \frac{G_F^2}{6\pi^2} m_{B_s} f_{B_s}^2 B_{B_s} \eta_B m_t^2 f_2(m_t^2/M_W^2)|V_{ts}^* V_{tb}|^2
\end{aligned}
\tag{24}
$$

and

$$
\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s} f_{B_s}^2 B_{B_s}}{m_{B_d} f_{B_d}^2 B_{B_d}} \left| \frac{V_{ts}}{V_{td}} \right|^2.
\tag{25}
$$

Assuming unitarity of the CKM matrix, $|V_{tb}| \approx 1$. Since the top quark mass has now been measured, the only unknowns in these formulae are the magnitudes of the CKM elements $V_{td}$ and $V_{ts}$, the decay constants $f_B$ (which in principle can be measured) and the bag factors, $B_B$, which are theoretical constructs.

Estimating $f_B$ and $B_B$ for the $B^0$ and $B_s^0$, then, is essential for extraction of the CKM elements. Since these mesons differ only by their light quarks, it is reasonable to expect that model dependences in predictions of, for instance, $f_{B_d}$ and $f_{B_s}$ are significantly reduced in the ratio $f_{B_d}/f_{B_s}$. These issues are addressed the next section.

## 2.3   The B Decay Constant and Bag Parameter

The $B$ decay constant $f_B$ is necessary to relate the $B$ oscillation frequency to the CKM matrix element $V_{td}$, as described in Section 2. Prospects for measuring this decay constant directly are rather bleak. In principle, it can be measured just as $f_\pi$ is measured, using leptonic decays. The partial width of the decay $B \to l\nu_l$ is

$$
\Gamma(B^+ \to l^+ \nu_l) = \frac{G_F^2 |V_{ub}|^2}{8\pi} f_B^2 m_B m_l^2 \left( 1 - \frac{m_l^2}{m_B^2} \right)^2,
\tag{26}
$$

where $l = e, \mu, \tau$. The decays into $e$ and $\mu$ are helicity suppressed, leaving the experimentally challenging $B \to \tau\nu_\tau$ as the dominant leptonic decay. The branching ratio into this mode, however, is expected to be very small. Using the range for $V_{ud}$ of 0.002-0.005 [6] and $f_{B_d} = 200$ MeV (see below) results in a branching ratio between $1.8 \times 10^{-4}$ and $3 \times 10^{-5}$.

ALEPH [17], DELPHI [18] and L3 [19] have placed 90% CL upper limits on this branching ratio, of $1.8 \times 10^{-3}$, $1.1 \times 10^{-3}$ and $5.7 \times 10^{-4}$ respectively. Since the LEP data is played out, no further improvement is likely. CLEO [20] searches for a $B \to \tau\nu_\tau$ signal by reconstructing the mass of the $B$ meson, using the beam constraint to account for the missing neutrinos. The background to this analysis is very high – the signal is $-9 \pm 36$ events. The CLEO group reports a 90% CL limit of $2.2 \times 10^{-3}$ based on a sample of $2.2 \times 10^6$ $B$ decays.

The most promising method to extract $f_B$ from experimental data had seemed to be via the $B^+ - B^0$ lifetime difference. It was recently estimated [21] that the lifetime ratio is

$$
\frac{\tau^-}{\tau^0} = 1 + 0.05 \frac{f_B^2}{(200 \text{ MeV})^2},
\tag{27}
$$

with an extra $\mathcal{O}(15\%)$ error added to account for terms of order $m_b^4$ that are not included in the calculations. Thus, a measurement of the lifetime ratio to 1%

corresponds to a measurement of $f_B$ to $\approx 20$ MeV, with a 30 MeV theoretical error. The current value of this ratio is $1.04 \pm 0.04$ [6]. CDF has measured it as $1.06 \pm 0.07 \pm 0.01$ in exclusive $B \to J/\Psi K$ decays [22] and as $1.11 \pm 0.06 \pm 0.03$ in semileptonic decays [23]. The $J/\Psi$ sample in CDF's upcoming Run II is expected to be more than 20 times larger than the current one, with a similar increase in the inclusive sample. Along with modest improvements in the systematic errors, this leads to an expected precision on this ratio of $\approx 1\%$. The BaBar group expects to measure this ratio to close to 1% as well [24]. Recently, however, our ability to predict lifetimes and lifetime ratios has been called into question because none of the models account for the short $\Lambda_b$ lifetime. This failure has led some authors [25] to suggest that some of the approximations used in deriving Equation 27 are not valid – specifically there are large nonperturbative effects that should not be ignored. Until the situation is clarified, extracting physical quantities from heavy quark lifetimes is suspect.

This leaves us with two ways to obtain $f_B$: lattice QCD and QCD sum rules.

In the past year or two, lattice calculations of $f_B$ have matured considerably. Our understanding of lattice discretation effects has grown, and the advent of nonrelativistic actions have provided a means to greatly reduce the uncertainty introduced by the relatively large lattice spacing.

One of the main problems with calculating $f_B$ on the lattice is that the $b$ quark is very heavy, and the wavelength of the $b$ quark is small compared to the lattice spacings currently computationally feasible. Typical lattice spacings, $a$, available today are $\mathcal{O}(0.1 - 0.2)$ fm, so one has $m_Q a \simeq 25a > 1$ (where $a$ is in fm). Thus, in extrapolating to the continuum limit, $m_Q a = 0$, one has a long way to extrapolate; further, one must contend with lattice-size artifacts which are $\mathcal{O}(m_Q a)$, $\mathcal{O}[(m_Q a)^2]$ and so on, and each may be sizeable.

The original Wilson lattice action was developed for light quarks, for which $ma \ll 1$ holds. This discretization procedure introduces lattice spacing artifacts of $\mathcal{O}(a)$, $\mathcal{O}(a^2)$ and so on, which makes the extrapolation to the continuum, $a = 0$, a major source of systematic uncertainty. The Sheikholeslami-Wohlert (SW or clover) action [26] was introduced as an improvement on the Wilson action for light quarks; it removes the leading $\mathcal{O}(a)$ lattice effects, resulting in a much-improved behaviour. Lattice calculations continue to use both of these actions for the light quarks in the $B$ and $D$ meson systems, along with a variety of approaches in handling the heavy quark.

In placing a heavy-light system on the lattice, different techniques were developed for handling the heavy quark. Taking the straight Wilson action is problematic because $m_Q a > 1$. One common method of dealing with the $b$ quark, known as the static approximation, is to set its mass to be infinite. In the first six or so years of lattice calculations this method was used extensively, and it produced results for $f_B$ which varied over an extremely wide range. By 1994, these variations were understood to be due to large contamination from excited $B^*$ states in the lattice version of the $B$ meson. Once this was corrected for, the results from simulations which used static $b$ quarks stabilized, with values comparable to those from other methods.

In the meantime, propagating $b$ quarks were introduced by performing the lattice calculations with finite quark masses, but closer to $m_c$, where $m_Q a < 1$. Typically, several calculations with different masses were performed, and then used to extrapolate up to $m_Q = m_b$ using a linear extrapolation in $1/m_Q$. Often the static limit was included to help anchor this process. In these methods,

systematic errors are introduced both by finite lattice spacing artifacts and the uncertainties in the mass extrapolation.

More recently the heavy $b$-quark problem has been addressed by using an effective nonrelativistic action for the $b$ quark. In the rest frame of a $B$ meson, the light valence quark has a momentum $\mathcal{O}(\Lambda_{\mathrm{QCD}})$. By momentum conservation, the heavy $b$ quark must also have momentum of this order. If the $b$ quark is nonrelativistic, $p = m_Q v = \Lambda_{\mathrm{QCD}}$, and so the velocity $v \approx \Lambda_{\mathrm{QCD}}/m_Q \approx 0.05$, justifying the nonrelativistic approximation. Thus, one can effectively expand quantities in powers of $1/m_Q$ and expect reasonable behaviour. Further, in a nonrelativistic effective theory, the rest mass of the $b$ quark does not appear, so we sidestep the problem of having a large $m_Q a$. Since the first use of this approach in 1994 [41], which included the $\mathcal{O}(1/m_Q)$ operators, various groups have continued to add more and more of the $\mathcal{O}(1/m_Q^2)$ terms.

In such nonrelativistic QCD (NRQCD) calculations, $1/m_Q$ effects and lattice spacing $a$ effects end up intertwined, so one of the challenges of such work is to effectively evaluate systematics due to the extrapolations. The uncertainties are much reduced if the next-order $1/m_Q^2$ terms are included. In early 1998 the MILC collaboration presented results with a comprehensive treatment of both these effects, with improvements through $\mathcal{O}[(\Lambda_{\mathrm{QCD}}/m_Q)^2]$ and $\mathcal{O}(a)$ [46].

A third technique, developed by the Fermilab group (KLM) [54], allows the use of either Wilson or clover actions, with suitable normalisation and nonrelativistic interpretation, for the quarks at any $m_Q a$. Simulations based on this heavy-quark action are run directly at $m = m_b$, so no mass extrapolation is required.

Until recently, all calculations of $f_B$ have been performed within the quenched approximation, i.e. no virtual quark loops were simulated. Estimating the size of the shift in $f_B$ due to this omission has been one of the great outstanding issues for the lattice community. There are now three groups reporting partially unquenched simulations, with two flavors of light fermions: the MILC collaboration [50, 55], Collins et al [51], and CP-PACS [52, 53]. The results are summarized in Table 3. All three groups found that unquenching the lattice raised all the $B$ and $D$ decay constants by 15-20%, although the results are still preliminary.

Table 3: Results for the decay constant $f_{B_d}$ from several lattice simulations with two dynamical sea quarks.

| Reference | $f_{B_d}$ (MeV) |
|---|---|
| MILC 99 [50] | $194 \pm 22^{+20}_{-0}$ |
| CP-PACS 99 [52, 53] | $210^{+9}_{-16} \pm 20$ |
| Collins 99 [51] | $186 \pm 25^{+50}_{-0}$ |
| Average | $200 \pm 30$ |

Figure 4 shows lattice results for $f_{B_d}$ over the 10-year history of the field. The results are grouped such that similar methods of dealing with the $b$ quark (static, mass extrapolation, KLM, NRQCD) are shown together for easier comparison. The points are grouped into three rough time periods: pre-1994 years, the early years; 1994-1997, during which NRQCD techniques were developed to include higher-order terms; and post-1997, when fully $\mathcal{O}(1/m_Q^2)$, $\mathcal{O}(a)$ improved simulations were implemented.
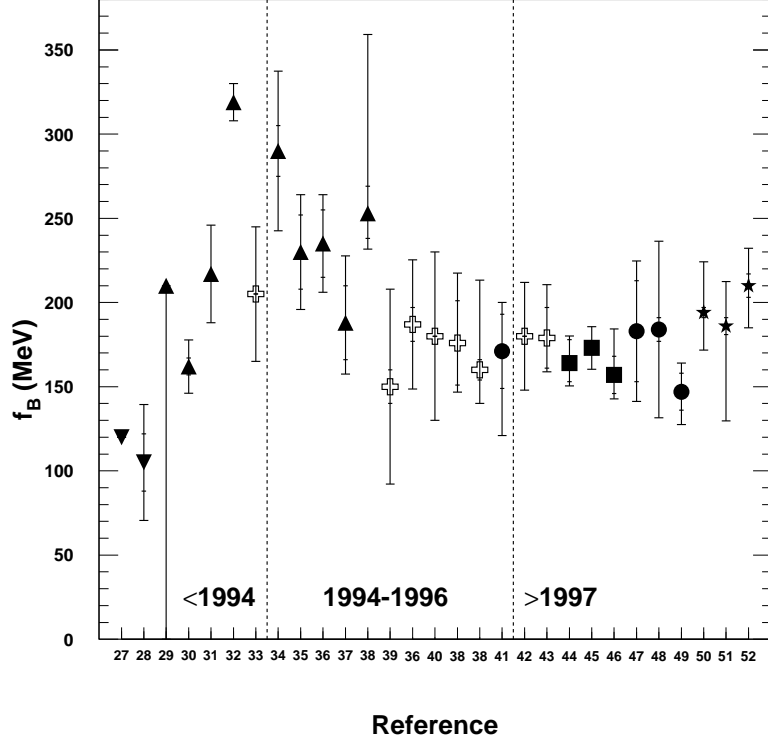
Figure 4: Summary of lattice calculations of the decay constant $f_{B_d}$. The results are grouped by the treatment of the heavy quark: down-pointing triangles, mass extrapolation from $m < m_c$; up-pointing triangles, static calculations; open crosses, mass extraploation from $m > m_c$ to $m_b$; circles, NRQCD; squares, the KLM action at $m_b$. The rightmost three results, marked with a star, are those with two dynamical fermions.

Table 4 shows averages of $f_{B_d}$, $f_{D_s}$, and $f_{B_s}/f_{B_d}$, using data from the collaborations shown. Results with full estimations of systematics were included in the fit. The average result, $f_{B_d} = 166 \pm 8$ MeV, does not include a systematic due to quenching. This is much lower than the average obtained from the $N_f = 2$ calculations, which is $200 \pm 30$ MeV (allowing a large systematic as these are still preliminary).

Unfortunately, checking the lattice and sum rule results against experiment is difficult, because experimental data are in short supply as noted earlier. However, there is one decay constant prediction which can be correlated with data, namely $f_{D_s}$. The branching ratio of $D_s \to \tau\nu_\tau$ is fairly large, $7 \pm 4\%$ [6], owing to the relative lack of hadronic decay modes in contrast to both $B$ and $D$ decays, which have ample phase space available for a multitude of hadronic modes. There are currently eight measurements [56] of the branching ratio $\mathrm{Br}(D_s \to \tau\nu_\tau)$, which, along with $V_{cs}$, allow us to extract $f_{D_s}$ as per Equation 26. Figure 5 summarizes these measurements.

Table 4: Lattice results from simulations selected as described in the text. A common systematic error due to quenching is not included.

| Reference | $f_{B_d}$ (MeV) | $f_{D_s}$ (MeV) | $f_{B_s}/f_{B_d}$ |
|---|---|---|---|
| MILC 98 [46] | $157^{+27}_{-14}$ | $210^{+27}_{-13}$ | $1.11 \pm 0.05$ |
| JLQCD 98 [45] | $173 \pm 13$ | $224 \pm 19$ | $1.18 \pm 0.06[63]$ |
| APE 98 [43] | $179^{+32}_{-20}$ | $231 \pm 13$ | $1.14 \pm 0.03$ |
| GLOK 98 [49] | $147 \pm 19$ | NA | $1.20 \pm 0.04$ |
| FNAL 98 [44] | $164 \pm 16$ | $213 \pm 18$ | $1.13 \pm 0.05$ |
| APE 97 [42] | $180 \pm 32$ | $237 \pm 16$ | $1.14 \pm 0.08$ |
| Average | $166 \pm 8$ | $226 \pm 8$ | $1.15 \pm 0.02 \pm 0.06^{\dagger}$ |

$^{\dagger}$ Following a recent review [58] a common systematic of 0.06 is assigned.

The ARGUS result is model-dependent and omitted from the average. Results from WA75 and E653 have been adjusted to take into account updated $D_s$ and $D^0$ branching ratios. Note that the errors have actually increased as the $D_s \to \phi\pi$ branching ratio measurement was downgraded by the PDG from having an error of 0.4% in 1994 to 0.9% in 1996. The measurements were combined, taking into account correlations in the branching ratios used to derive $f_{D_s}$. The result is

$$f_{D_s}^{\text{exp}} = 273 \pm 28 \quad \text{MeV}. \tag{28}$$

We obtain a lattice average of $f_{D_s} = 226 \pm 8$ MeV, not including quenching errors, from the same group of collaborations used to derive $f_{B_d}$, as shown in Table 4. There is reasonable agreement between the measurement and the lattice result, though the agreement would be better using the unquenched lattices. More precise experimental data and a better understanding of the unquenched lattices would be of great help.

Since $f_{D_s}$ is likely to be the only experimental input for many years, an interesting exersize is to bootstrap $f_{B_d}$ and $f_{B_s}$ from it. The idea is that many of the largest systematic errors in the lattice calculation of the heavy-light decay constants either cancel, or are largely ameliorated, when one takes a ratio. For example, unquenching the lattice is expected to have a similar effect on $f_{B_d}$ and $f_{D_s}$, as should lattice size corrections. Several groups [43, 46] are now calculating the ratio $f_{B_d}/f_{D_s}$ directly from the same simulation. Multiplying the average value of $f_{B_d}/f_{D_s} = 0.78^{+6}_{-4}$ by the experimental value for $f_{D_s}$ results in:

$$f_{B_d} = \frac{f_{B_d}^{\text{lat}}}{f_{D_s}^{\text{lat}}} f_{D_s}^{\text{exp}} = 213 \pm 22(\text{expt.})^{+16}_{-11}(\text{theor.}) \tag{29}$$

The other technique used to calculate $f_B$ is based on QCD sum rules. (For an excellent summary of this method, see Reference [57].) The sum rules are derived from the two-point correlation function:

$$\Pi(q^2) = i \int d^4 x e^{iqx} < 0|T\bar{q}(x)i\gamma_5 b(x), \bar{b}(0)i\gamma_5 q(0)|0 > .$$

Inserting a complete set of states with $B$-meson quantum numbers in the time-ordered product gives

$$\Pi(q^2) = \frac{< 0|\bar{q}i\gamma_5 b|B >< B|\bar{b}i\gamma_5 q|0 >}{m_B^2 - q^2} + \sum_h \frac{< 0|\bar{q}i\gamma_5 b|h >< h|\bar{b}i\gamma_5 q|0 >}{m_h^2 - q^2}.$$
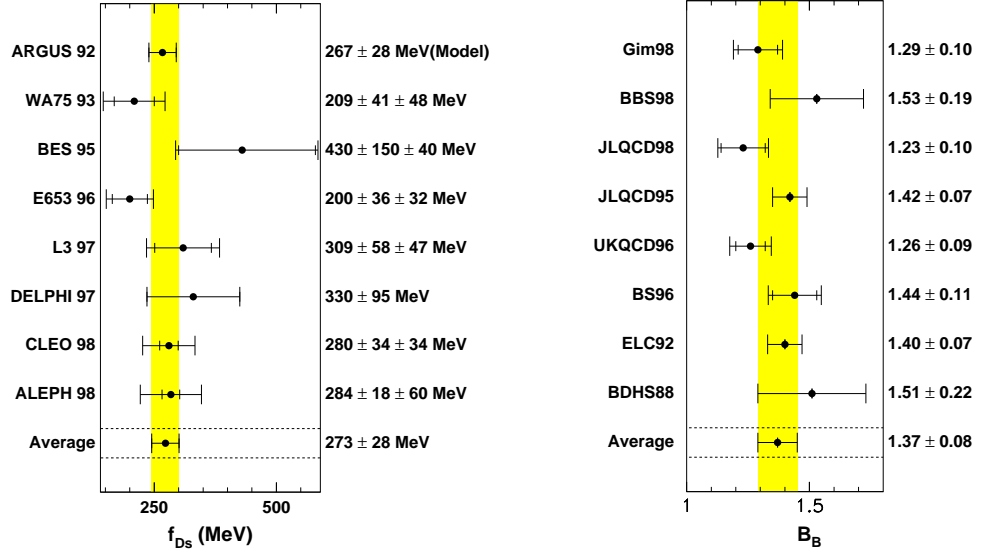
Figure 5: *Left:* Average of $f_{D_s}$ experimental measurements. The model-dependent ARGUS result is not included in the average. *Right:* Average of lattice calculations of $\hat{B}_{B_d}^{\text{NLO}}$. These have been adjusted so that the NLO corrections are handled in a uniform manner, and include a common systematic of 0.07 due to uncertainties in the renormalization constant.

Using $m_b < 0|\bar{q}i\gamma_5 b|B \rangle = m_B^2 f_B$, we see that the first term contains $f_B^2$.

The two-point correlation function can also be expanded using an operator product expansion (OPE):

$$\Pi(q^2) = \sum_d C_d(q^2, \mu)\hat{O}_d(\mu), \tag{30}$$

where $C_d$ are the (calculable) Wilson coefficients of the operators of various dimension $d$ which appear in the expansion.

Equating these expressions for $\Pi(q^2)$, along with experimental input on the expectation values of the operators $\hat{O}_d$ (e.g. from bottomonium and charmonium masses), allows the extraction of $f_B$.

The OPE expansion of $\Pi(q^2)$ is rather sensitive to the value of the $b$-quark mass chosen, which ends up being the dominant systematic effect. For example, taking $m_b = 4.7 \pm 0.1$ GeV results in $f_B = 180 \pm 30$ MeV [57].

Given that unquenching the lattice seems to shift $f_{B_d}$ by 15-20% and that several groups confirm this result, we choose to use the preliminary results with $N_f = 2$ rather than perpetuate what seems to be a low $f_{B_d}$ value. The quenched results on $f_{B_d}$ measured in a multitude of ways seem to have stabilized, which suggests that the unquenched simulations will soon have errors under much better control.

Combining the three values:

$$f_{B_d} = 200 \pm 30, \qquad\qquad\qquad \text{(Lattice, } N_f = 2\text{)}$$

$$f_{B_d} = 180 \pm 30, \qquad \text{(QCD sum rules)}$$
$$f_{B_d} = 213 \pm 22(\text{expt.})^{+16}_{-11}(\text{theor.}) \qquad (f^{\text{lat}}_{B_d}/f^{\text{lat}}_{D_s} \times f^{\text{exp}}_{D_s})$$

gives our estimate: $f_{B_d} = 200 \pm 16$ MeV.

### 2.3.1 THE BAG PARAMETER

The second quantity standing between measurements of $\Delta m_d$ and the clean extraction of $V_{td}$ is the so-called bag parameter $B_B$. $B_B$ represents a "fudge-factor" introduced because we collapsed the sum over all intermediate states in Equation 20 to just the vacuum intermediate state. If the vacuum is the only state which ends up contributing, it saturates the sum, and $B_B = 1$.

As such, $B_B$ is a completely theoretical construct and cannot be measured experimentally (only the product $f_B^2 B_B$ has physical meaning). $B_B$ is also amenable to study on the lattice, and many estimates of its value have been made by the same groups which calculate $f_B$.

To calculate $B_B$, the ratio

$$r = \frac{<\overline{B}|\hat{O}_{LL}|B>}{<\overline{B}|\hat{O}_{LL}|B>_{\text{Vac}}} = \frac{<\overline{B}|\hat{O}_{LL}|B>}{8/3 f_B^2 m_B^2}$$

is evaluated using standard lattice techniques (where Vac refers to the vacuum saturation approximation). As with $f_B$, various groups have used static $b$ quarks, propagating $b$ quarks with mass extrapolation, and NRQCD actions. $B_B$ depends on the choice of scale $\mu$, and groups quote their results at scales ranging from 2 GeV to 5 GeV. The lattice and continuum operators must then be matched at this scale. The physical quantity of interest – the NLO, renormalization-group-invariant quantity, is

$$\hat{B}_B^{\text{NLO}}(\mu) = \alpha_s(\mu)^{-2/\beta_0}(1 + \frac{\alpha_s(\mu)}{4\pi}J)B_B(\mu), \tag{31}$$

where
$$
\begin{array}{llll}
J & = & \frac{\gamma_0 \beta_1}{2\beta_0^2} - \frac{\gamma_1}{2\beta_0}, & \gamma_0 & = & 4 \\
\beta_0 & = & 11 - \frac{2n_f}{3}, & \gamma_1 & = & -7 + \frac{4n_f}{9} \\
\beta_1 & = & 51 - \frac{19n_f}{3}
\end{array}
$$

Throughout the literature, a number of different choices are present, e.g. $n_f = 0, 4, 5$, $\Lambda^{(4)}_{\text{QCD}} = 200, 239$ MeV, $\Lambda^{(5)}_{\text{QCD}} = 183, 200$ MeV, $m_b = 4.33, 5.0$ GeV, as well as scaling with $J = 0$.

Following Reference [59], we have tried to rescale the results given in a consistent manner; namely, $n_f = 4$, $\Lambda^{(4)}_{\text{QCD}} = 200$ MeV, and $\mu = 5$ GeV are used to scale $\alpha_s$ up to the matching point at 5 GeV, with $\alpha_s$ given by the NLO expression [6]. We match $\alpha_s$ between $n_f = 4$ and $n_f = 5$ at the $b$ mass of 5 GeV, and set $n_f = 5$ in the exponent of $\alpha_s$, $-2/\beta_0$. This results in a scaling of about 1.6 between the raw $B_B(m_b)$ and $\hat{B}_B^{\text{NLO}}(m_b)$.

The results on $B_{B_d}$ from calculations using static $b$ quarks were omitted from recent averages due to an extremely large variation in the results when the NLO corrections were included in three supposedly equivalent ways [60, 61]. Recently, several clarifications to this process were made by Gimenez & Reyes [59]. First, the calculation of the lattice operator $\hat{O}_{LL} = \overline{b}\gamma^\mu(1 - \gamma_5)q\overline{b}\gamma_\mu(1 - \gamma_5)q$ for SW (clover) actions used previously was incorrect, except for those groups using

Wilson actions for the light quarks. Second, the importance of using tadpole-improved light quarks was shown. Third, some inconsistencies in the techniques used to match operators to the continuum were pointed out. They have reanalyzed their data [61] along with data from the UKQCD group [65], which also used SW light quarks. The results are now much more stable under variations in the matching scheme, as well as being in good agreement with the data from simulations that used propagating heavy quarks [33, 64, 66].

Figure 5 shows the average for $\hat{B}_{B_d}^{\mathrm{NLO}}$ from the analyses in References [28, 33, 59, 62, 63, 64, 65, 66], which is $\hat{B}_{B_d}^{\mathrm{NLO}} = 1.37 \pm 0.08$. A common systematic error of 0.07, due to uncertainties in the renormalization, has been included.

The ratio $B_{B_d}/B_{B_s}$ is expected to be very close to one, and has been estimated as $B_{B_d}/B_{B_s} = 1.01 \pm 0.01$ [61]

## 3    ANATOMY OF A TIME-DEPENDENT OSCILLATION MEASUREMENT

We now turn to the experimental side of explicitly measuring the time dependence of the $B\overline{B}$ flavor oscillations. This section summarizes the basic analysis elements and techniques that have been used to produce the results presented in the following two sections.

The basic steps of any time-dependent mixing measurements are evidently as follows:

1. Find the interaction point at which the $B$ mesons were produced (primary vertex) and decay point of one of the mesons (secondary vertex), and determine the decay distance $L$.

2. Measure the (presumed) $B$ meson's momentum to turn the decay distance $L$ into the proper time lived, $t$.

3. Tag the flavor of the (presumed) $B$ at its production point, and determine the probability the tag was correct, $P_r^P(t)$.

4. Tag the flavor of the $B$ at its decay point, and determine the probability the tag was correct, $P_r^D(t)$.

The advent of high-precision silicon tracking systems allows the measurement of the decay distances of $B$ mesons, which are typically of order $\gamma\beta c\tau_B \approx \gamma\beta \times 470\mu$m. With boost factors ranging from $\beta\gamma \sim 3$ at CDF to $\beta\gamma \sim 7$ at LEP and SLD, decay lengths in the millimeter range are typical, which are easily discerned by modern detectors. For detectors with three-dimensional precision tracking (i.e. double-sided silicon or pixel detectors) $t = L/\gamma\beta c = LM_B/p_B$. Experiments with single-sided silicon strip detectors measure the flight distance projected onto the plane perpendicular to the beamline ($xy$, or $r\phi$ plane), $L_{xy}$, and convert to $L$ using the estimated polar angle of the meson, $t = L_{xy}/(\gamma\beta c\sin\theta) = L_{xy}M_B/p_t$.

The uncertainty on the decay time of the meson in its rest frame is

$$\sigma^2(t) \;\; = \;\; \frac{\sigma^2(L)}{\gamma^2\beta^2 c^2} + t^2\frac{\sigma^2(\gamma\beta)}{\gamma^2\beta^2}. \tag{32}$$

The resolution on $L$ (or $L_{xy}$) is typically $\mathcal{O}(300\mu$m) at LEP and $\mathcal{O}(100\mu$m) at CDF, so when divided by $\gamma$ does not contribute significantly to the error on $ct$ compared to the momentum-resolution term $\sigma(\gamma\beta)$, which grows linearly
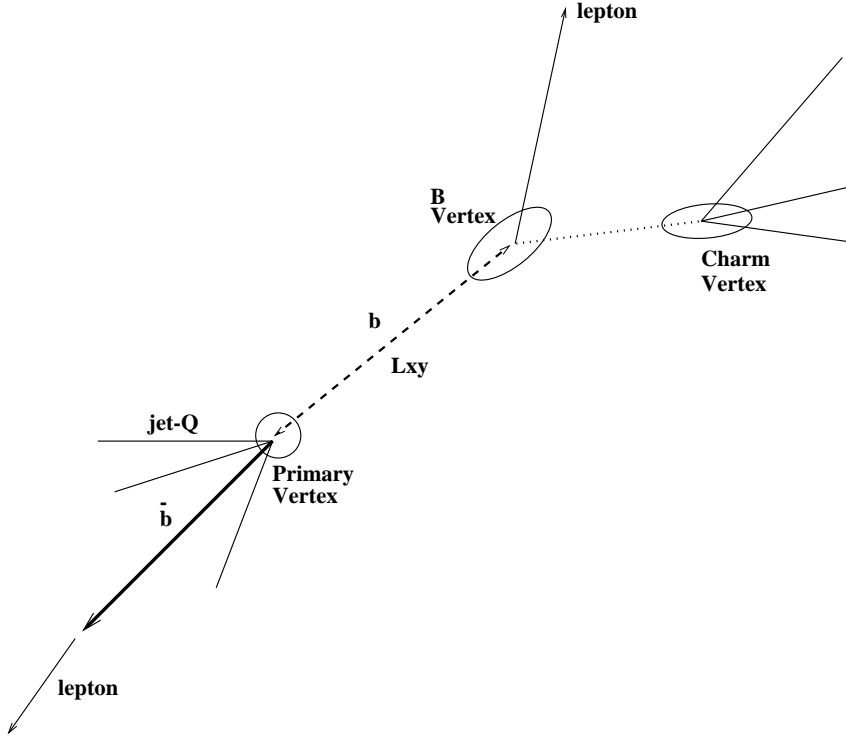
Figure 6: Schematic of a typical $B$ event.

with the decay length $L$. This is a critical point for measuring $B_s^0$ oscillations, which have a very rapid oscillation rate. Once the smearing induced by the momentum resolution reaches half the oscillation period, any hint of an oscillation signal is washed out. With the momentum resolution achievable from partially reconstructed decays, only the first few oscillations of the $B_s^0$ will be visible, making frequency measurement very difficult. Fully reconstructed $B_s^0$ decays, of course, have optimal momentum resolution, but current experiments have very small data samples.

The variety of mixing measurements comes mostly from the wide range of techniques used to identify the $b$-quark flavor at its production and decay point. These techniques are breifly described below.

By far the dominant production mode of $b$-type quarks is $b\bar{b}$ pair production, from $e^+e^-$ annihilation at LEP and SLD, and from quark-antiquark annihilation and gluon-gluon fusion at CDF. At collider experiments they are generally produced roughly back-to-back, so the event can be divided into two hemispheres (defined by plane perpendicular to a vector $\hat{a}$ passing through the event origin) that each contain one of the $b$ quarks. The analyses are insensitive to slight changes in the direction of $\hat{a}$, and several choices for $\hat{a}$ are used, including: the thrust or sphericity axis of the event; the direction between the primary and secondary vertices; or the direction of a jet associated with the secondary vertex, defined by an appropriate clustering algorithm. The hemisphere with the secondary vertex is called the vertex side, and the other is called the opposite side. (This distinction becomes arbitrary if vertices are reconstructed in both the $b$ and $\bar{b}$ hemispheres.)

Figure 6 is a sketch of an event, showing the $b$ and $\overline{b}$ quarks originating from the primary vertex, a secondary vertex found on the $b$ (vertex) side, and some possible flavor tags on the opposite side (see below). The hemispheres would be formed by a plane passing through the primary vertex roughly perpendicular to the $b$-$\overline{b}$ axis.

## 3.1  Decay-Point Flavor Tagging

The $b$-quark flavor at decay is typically identified by one of two methods, either by the charge of the lepton($e$ or $\mu$) from a semileptonic decay of the meson, or by the type of $D$ meson present in a partial reconstruction of the decay.

Table 5: Sources of leptons from $b$ quark decay (correct flavor tags of the $b$ quark are denoted r=right, incorrect tags are denoted w=wrong).

| Decay chain | Lepton Source | Correct Tag? (r/w) |
|---|---|---|
| $b \to W^- \to l^-$, | Direct | r |
| $b \to c \to W^+ \to l^+$ | Cascade | w |
| $b \to W^- \to \overline{c} \to l^-$ | Right-sign cascade | r |
| $b \to W^- \to \tau^- \to l^-$ | Right-sign cascade | r |
| $b \to J/\psi X \to l^{\pm}$ | Both-sign cascade | r/w |

A lepton from a semileptonic $b$ decay can have the sources given in Table 5. The main contributions come from the first three components. The relative rates of leptons from direct and wrong-sign sequential decays are approximately equal [6] The right-sign sequentials $b \to \overline{c} \to l^-$ contribute approximately the same amount as the wrong-sign [6], so the lepton tags the $b$-quark correctly approximately 2/3 of the time. However, with very simple kinematic cuts, the contribution from the sequential decays can be greatly reduced. In particular, the momentum spectrum of the sequential leptons is softer than that of the direct leptons, as they arise further down the decay chain.

The second common method of tagging the decay-point $b$ charge is by either partially or fully reconstructing the charmed meson from the $B$ decay. In particular, the decays $\overline{B^0} \to D^{*+}X$, $(D^{*+} \to D^0\pi_*^+)$ and $\overline{B^0} \to D^+X$ allow the separation of $b$ from $\overline{b}$ quarks. ($\pi_*$ denotes the soft pion from the $D^*$ decay.) The sign of the $b$-quark charge is the same as the sign of the kaon charge in the $D$ meson decay, and it is opposite that of the $\pi_*$ from the $D^{*+}$ decay.

Figure 7 shows the decay modes of the $\overline{B^0}$ and $B^-$ assuming final-state interaction effects are small (i.e. factorization approximately holds). Soft gluons running around the diagrams are not shown. In the semileptonic modes (Figure 7$a,b$), note that the $B^-$ produces a $D^0, D^{*0}, D^{**0}$..., and that the $D^{*0}$ decays only to $D^0$, not $D^+$. Thus the only contamination of $B^-$ in the $D^{*+}$ sample is through decays involving $D^{**}$. In contrast, the $\overline{B^0}$ decay can produce a $D^+$ directly, or through a $D^{*+}$ with a missing $\pi^0$ and a $D^{*+}$ detected in the $D^0\pi_*^+$ configuration. The decays involving $D^{**}$ will end up scattered around the various modes, as in the $B^-$ case.

Hadronic decays show the same basic behavior, with the addition of the internal, color-suppressed diagrams (Figure 7$e,f$). Note that the type of charm decay
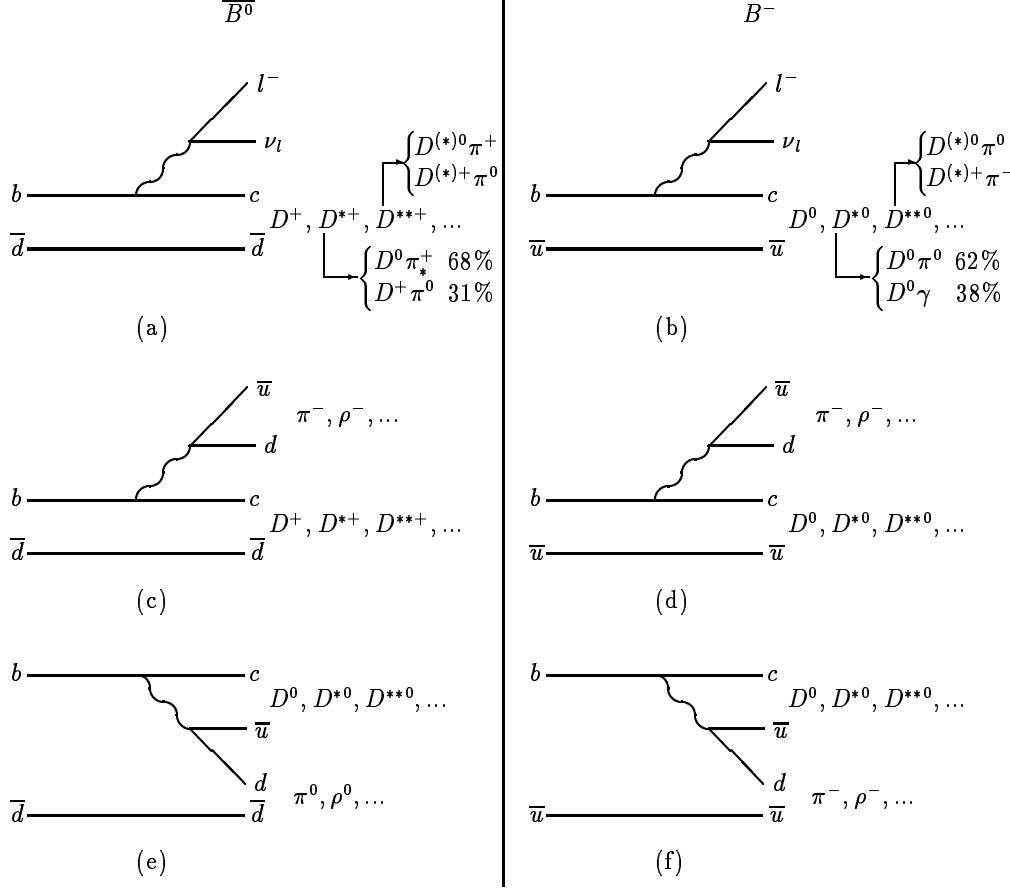
Figure 7: Diagrams for semileptonic and hadronic decay of $\overline{B^0}$ and $B^-$ mesons (soft gluons are suppressed).

product for these diagrams is the same for the $B^-$ decays. The $\overline{B^0}$, however, would tend to produce $D^0, D^{*0}, ...$, identically to the $B^-$. This breaks the analogy between the semileptonic and hadronic modes, resulting in a slightly different fraction of $D^{*+}$ coming from $\overline{B^0}$ in the two cases, but the effect is not large.

Diagrams like those in Figure 7c and 7d can also have the external $W$ decaying into a $\overline{c}s$ pair, producing a charm of the same charge as the $b$ quark. Fortunately, the branching ratio of $B \to D^{(*)} \overline{D}_s^{(*)}$ is only $\mathcal{O}(5\%)$.

Evidently, requiring a reconstructed $D^{*+}$ (with or without an associated lepton) in the final state not only flavor-tags the $b$ quark at the decay point but also greatly enriches the sample's $B^0$ fraction. Although no measurement of the fraction of $D^{*+}$ that comes from $\overline{B^0}$ (as opposed to $B^-$) is available for hadronic decays, it has been measured to be, e.g., 84±9% by OPAL [67] in the semileptonic mode.

The advantages to the $D^{*+}$ tag are its excellent purity and its ability to produce event samples with very high $B^0$ content. Its obvious drawback is its efficiency, which is rather low owing to the exclusive final state.

A third method of tagging the $b$-quark flavor at the decay point is via the charge of the kaon(s) in the decay chain. The decay of the $b$ quark proceeds $b \to c \to s$. Hence a $K^-$ tags a $b$ quark and a $K^+$ tags a $\overline{b}$. Using inclusive kaons will greatly increase the efficiency of this tagging method over the $D^*$ method, but of course there are other sources of kaons besides the $D$-meson decay, which will dilute the tag's effectiveness. In addition, the dramatic increase in the $B^0$ content of the sample is lost.

Finally, a method based on a charge dipole is used by SLD (see Section 4.6).

### 3.2   Production-Point Flavor Tagging

The flavor tagging of the $b$ quark is more problematic at its production point than at its decay point. Broadly speaking, there are two strategies: same-side techniques, which use information from the vertex hemisphere to determine the initial flavor of its $b$ quark, and opposite-side techniques, which measure the charge of the $b$ quark in the opposite hemisphere (at either its decay or production point) and infer the vertex-side $b$ charge from that.

Opposite-side tags come in two types: lepton tags and jet charge. Lepton tagging is essentially the same as described above for decay-point flavor tagging, although the selection criteria on the lepton may differ for the two sides. This type of tagging has high purity but suffers from a low efficiency because the semileptonic branching fraction of $B$ mesons is only 20%; in addition, detection and selection criteria are imposed.

The jet-charge method relies on a weak correlation between the momentum and charge of tracks in a $b\overline{b}$ event. For example, at SLD or LEP, $b$ and $\overline{b}$ quarks are generally produced back-to-back each with momentum of 45 GeV. Thus the $b$ quark hemisphere starts off with a net $-1/3$ charge. As the fragmentation proceeds, the $b$ quark gradually loses its momentum as it radiates off gluons; however, the $-1/3$ charge continues to be associated with a higher-momentum particle, whereas the charge required to keep the hemisphere charge integral is in lower-momentum particles. Experimentally, a momentum-weighted charge distribution is used to form a variable to discriminate between $b$ and $\overline{b}$ quarks. Section 4.3 provides details on the algorithms used to measure the jet charge.

Two types of same-side tagging (SST) are commonly used. One method relies on an even more specific charge correlation than the jet charge [68]. Figure 8 shows a schematic of a $\overline{b}$-quark fragmentation chain. If the first quark-antiquark pair radiated is a $u\overline{u}$, the $\overline{b}$ becomes a $B^+$, and the first charged pion in the fragmentation chain is a $\pi^-$. If the first $q\overline{q}$ pair is a $d\overline{d}$, the meson is a $B^0$ and the first charged pion is a $\pi^+$. Hence for a $B^0/\overline{B^0}$ sample, a correlated pion $\pi^+$ tags a $B^0$ and $\pi^-$ tags a $\overline{B^0}$. The correspondence is opposite for charged $B^\pm$. This tagging technique, then, requires a rather clean sample of $B^0/\overline{B^0}$ events, which rules out applying it to purely leptonically tagged events. The difficultly lies in selecting the correct pion from all the fragmentation tracks. In the case of a fragmentation into a $\overline{B^0_s}$ meson, the first fragmentation track would be a $K^-$.

The second same-side method is unique to SLD, and relies on the large longitudinal polarization that the SLC was able to achieve. The differential cross section for producing a $b$ (as opposed to $\overline{b}$) quark in a polarized $e^+e^-$ interaction is

$$\sigma_b(\cos\theta) \equiv \frac{d\sigma_b(\cos\theta)}{d\cos\theta} \quad \propto \quad (1 - A_e P_e)(1 + \cos^2\theta) + 2A_b(A_e - P_e)\cos\theta,$$
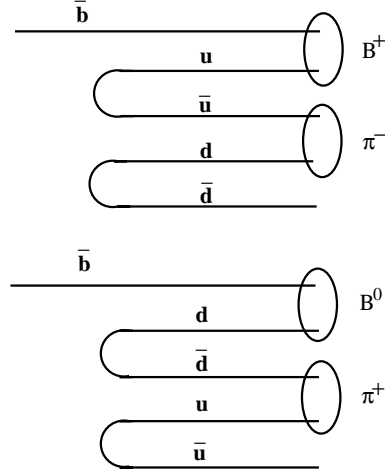
Figure 8: Schematic of $B^0$ and $B^+$ fragmentation chains

where $\theta$ is the angle between the $b$ quark and the incoming electron direction. $P_e$ is the longitudinal polarization of the electron beam ($P_e > 0$ for positive helicity, right-handed polarization), and $A_f = 2v_f a_f/(v_f^2 + a_f^2)$ where $a_f, v_f$ are the axial and vector coupling of the fermion $f$ to the $Z^0$. Thus the probability a $b$-type quark produced at angle $\theta$ is a $b$, rather than $\bar{b}$ is simply

$$\frac{\sigma_b(\cos\theta)}{\sigma_b(\cos\theta) + \sigma_{\bar{b}}(\cos\theta)} = \frac{1}{2} + \frac{A_b(A_e - P_e)\cos\theta}{(1 - A_e P_e)(1 + \cos^2\theta)} \equiv P_{\text{pol}}(\cos\theta)$$

(Some papers have the sign of $P_e$ swapped and/or measure $\theta$ from the positron direction.) The polarization of the electron beam is randomly chosen between the positive and negative helicity states on a pulse-by-pulse basis, and the signed polarization is recorded in the data stream. The standard model values for the asymmetries are $A_e = 0.155$ and $A_b = 0.94$ [6]. Note that the discriminating power that survives at $P_e = 0$ because of the residual forward-backward $b$ asymmetry has a maximum amplitude of $A_e A_b \sim 0.1$, whereas for large $P_e$ this amplitude is approximately $P_e A_b \sim P_e$. Hence SLD can tag the flavor of the $b$ quark at production simply by measuring the direction of the $B$-meson system. The probability that the sign assigned is correct, for a given angle, is simply the $P_{\text{pol}}(\cos\theta)$ above. It is the large (63 and 77%) polarizations achieved at the SLC that make this flavor tag effective, with a mean right-tag probability $<P_{\text{pol}}>$ of 62% and 76% respectively. Figure 10 shows a typical distribution of $\cos\theta$, along with the underlying predictions for $b$ and $\bar{b}$ quarks given by the Monte Carlo simulations.

Finally, many experiments combine multiple tags, both same-side and opposite-side, to maximize the probability of tagging correctly. For example, the SLD group uses the event-by-event $P_{\text{pol}}$ in all their analyses, along with a second flavor-tagging technique (e.g. jet charge). The two tags, $P_1$ and $P_{\text{pol}}$, are combined into the final right-tag probability

$$P_{\text{comb}} = \frac{P_1 P_{\text{pol}}}{P_1 P_{\text{pol}} + (1 - P_1)(1 - P_{\text{pol}})}. \tag{33}$$

### 3.3   Fitting Procedure

The signature for $B$-meson mixing is that the sign of the $b$ quark inferred at the decay point, $D$, is opposite to that inferred at production, $P$. In the following, the notation $DP+$ refers to events in which the experimental flavor-tag signs for the vertex-side $b$ quark at production and decay are the same, (i.e. the event looks unmixed) and $DP-$ is used when they are opposite (i.e. the event looks mixed). For example, in an analysis that uses leptons to tag both the production and decay charges, $DP+$ events are those with opposite-sign leptons. The event sample is then split into two components, with differing time dependences $N_{DP+}(t)$ and $N_{DP-}(t)$.

The basic variable to be measured, then, is the time-dependence of the decay and production flavor tags' charge correlations, $q_c(t) = \text{sign}(DP)(t)$. The mean value of this variable, binned in time $t$, is the standard measured asymmetry:

$$< q_c(t) > \equiv \mathcal{A}_m(t) \quad = \quad \frac{N_{DP+}(t) - N_{DP-}(t)}{N_{DP-}(t) + N_{DP-}(t)}.$$

Let us first consider this time-dependent charge correlation on a sample of pure $B^0/\overline{B^0}$. The numbers of events in the $DP+$ and $DP-$ samples at a given time are

$$N_{DP+}(t) \quad = \quad N \left( P_u(t)[P_r^P P_r^D + (1 - P_r^P)(1 - P_r^D)] \right.$$
$$\left. + \quad P_m(t)[P_r^P(1 - P_r^D) + (1 - P_r^P)P_r^D] \right)$$
$$N_{DP-}(t) \quad = \quad N \left( P_m(t)[P_r^P P_r^D + (1 - P_r^P)(1 - P_r^D)] \right.$$
$$\left. + \quad P_u(t)[P_r^P(1 - P_r^D) + (1 - P_r^P)P_r^D] \right),$$

where $N$ is the total number of events in the sample, $P_u$ and $P_m$ are the unmixed and mixed probabilities from Equation 14, and $P_r^P(P_r^D)$ is the probability that the production (decay) flavor tag got the sign of the $b$ quark correct. The asymmetry is

$$\mathcal{A}_m(t) \quad = \quad (2P_r^P - 1)(2P_r^D - 1)\frac{P_u(t) - P_m(t)}{P_u(t) + P_m(t)} = \mathcal{D}_P \mathcal{D}_D \cos \Delta mt.$$

That is, the measured asymmetry $A_m$ is reduced from the true asymmetry $A(t) = \cos \Delta mt$ by the production and decay tag dilutions, $\mathcal{D}_P = 2P_r^P - 1$ and $\mathcal{D}_D = 2P_r^D - 1$.

Now suppose we have a variable $\alpha$, calculated event-by-event, on which the production tag probability depends. If we break up the sample into bins of $\alpha$, with $N(\alpha)$ events in a given bin, then (setting $P_r^D = 1$ to reduce algebraic clutter) we have

$$N_{DP+}(t, \alpha) - N_{DP-}(t, \alpha) \quad = \quad N(\alpha)(2P_r^P(\alpha) - 1)\left[ P_u(t) - P_m(t) \right]$$

The error on this difference is $\sqrt{N(\alpha)}$, independent of the dilution $\mathcal{D}(\alpha)$, while the value of the difference scales with $\mathcal{D}(\alpha)$.

If we perform the time-dependent fit separately for each $\alpha$ bin, the relative errors on the fits will clearly scale as $\sigma(\alpha) \sim 1/[(2P_r(\alpha) - 1)\sqrt{N(\alpha)}]$. These (statistically) independent measurements are optimally combined using the weighted mean. This is equivalent to weighting each event in the sample by its dilution, $2P_r^P(\alpha) - 1$, and fitting once on this weighted sample. A similar result holds for the case with both $P_r^P$ and $P_r^D$ depending on a set of event-by-event variables $\vec{\alpha}$.

Thus in a time-binned fit to the asymmetry, the optimal procedure is to weight each event's contribution by the estimated dilution for that event. Most of the analyses described here use an unbinned likelihood fit, applied to the $N_{DP+}(t)$ and $N_{DP-}(t)$ samples simultaneously. In this case, employing the event-by-event right-tag probability $P_r^P(\alpha)$ is optimal.

The likelihood functions for the $DP+$ and $DP-$ subsamples, or the binned asymmetry, are constructed based on predicted proper time distributions of the data sample's components. These typically include unmixed $B^0$ decays, mixed $B^0$ decays, $B^+$, $b$ baryons, unmixed $B_s^0$, mixed $B_s^0$, cascade decays from all the above $b$ sources, charm background, and fake or combinatorial background.

The measured decay time in an event differs from the true decay time because of imperfect reconstruction of the decay length and of the $B$ momentum, as shown in Equation 32. The probability of reconstructing a decay time $t$ given that the true decay time was $t'$ is described by the resolution function $R(t, t')$. This is convoluted with the predicted true distribution, e.g., $P_u(t)$ or an exponential, to get the expected measured distribution. The vertexing algorithm might also introduce a $t$-dependent efficiency (actually $L$-dependent, which roughly translates into $t$-dependent), which is sometimes accounted for implicitly in $R$, or by multiplying the above convolution with an efficiency curve.

$R$ is typically parameterized as a sum of two or three gaussians, with separate parameterizations over several bins of decay time to allow for the increased smearing at longer times. If the vertex efficiency is included, the gaussians can differ above and below the mean to fit, e.g., a decrease in the probability to reconstruct a vertex near the primary vertex.

Alternatively, some analyses fit directly to the predicted distribution as a function of decay length, in which case the resolution function is parameterized as a function of decay length rather than time.

# 4 $B^0$ MIXING

The time-dependence of $B^0 - \overline{B^0}$ oscillations was first observed by ALEPH [69]. Since then, a number of analyses have measured the oscillation frequency, as summarized in Table 6. This section presents an overview of these measurements grouped by the techniques used to flavor tag the initial and final charges of the $b$-quark. Several preliminary results have been included, which are indicated on the summary in Figure 12.

## 4.1 Detector Overviews

What follows is a brief overview of the six collaborations' detectors, focussing on the elements which are most relevant to mixing analyses. More detailed descriptions are available for ALEPH [70], CDF [71], DELPHI [72], L3 [73], OPAL [74] and SLD [75].

All of the detectors operate at colliders and have similar cylindrical geometries with some type of end plug to cover most of the solid angle. Five of the experiments (ALEPH, DELPHI, L3, OPAL at LEP and SLD at SLC) operate at $e^+e^-$ machines running at the $Z$-pole, $\sqrt{s} = 91$ GeV. CDF operates at the Fermilab $\overline{p}p$ collider, which has an energy $\sqrt{s} = 1.8$ TeV.

Time-dependent oscillation measurements clearly require precision tracking detectors, and it was only after the inclusion of various silicon strip and pixel devices

Table 6: Index of the various flavor-tagging combinations used by each experiment to measure $B^0$ mixing. See the references indicated for details on individual analyses.

| Vertex tag | $l$ | | $D^{(*)}$ | | $D^{(*)}l$ | | | JetQ | $K$/dipole |
|---|---|---|---|---|---|---|---|---|---|
| Prod. Tag | $l$ | JetQ | $l$ | JetQ | $l$ | JetQ | SST | JetQ | JetQ |
| ALEPH | [76] | [76] | [76] | [76] | | | | [89] | |
| CDF | [77],[78],[79] | [78] | [86] | | [88] | | [91] | | |
| DELPHI | [80] | [80] | [80] | | [80] | | | | |
| L3 | [81] | [81] | | | | | | | |
| OPAL | [82] | [83] | [87] | | | [87] | | | |
| SLD | | [84],[85] | | | | | | | [90] |

that each experiment was able to perform these analyses. ALEPH, DELPHI and OPAL had silicon detectors in place since 1991. The ALEPH detector measures both the $r\phi$ and $rz$ views – OPAL(1993), DELPHI(1994) and L3(1994) soon added vertex detectors which could do the same. SLD had a CCD pixel device available since 1994.

Throughout the following, $d$ denotes the impact parameter, or distance of closest approach, of a track to the beam direction. The momentum $p$ of a track projected onto the $r\phi$ plane is denoted $p_t$.

## 4.1.1   PRIMARY VERTEX RECONSTRUCTION

The position of the interaction point of an event (primary vertex) is an important part of determining the decay length of the $b$ particle. The dimensions of the beam cross section (i.e. in the $x - y$ plane) are small for all the accelerators used in these analyses: $150\mu$m $\times$ $10\mu$m at LEP, $25\mu$m $\times$ $25\mu$m at Fermilab, and $1\mu$m $\times$ $1\mu$m at the SLC. Each experiment tracks the mean position of the beam centroid, giving an accurate estimate of the interaction point. In addition, some experiments fit for a vertex event-by-event, combining this mean position with tracks which have a high probability of originating there. This helps reduce the relatively large uncertainty on the $x$ position of the primary vertex at the LEP experiments.

## 4.1.2   PARTICLE IDENTIFICATION

The analyses presented here rely heavily on detecting and tagging $B$ mesons through their semileptonic decays to electrons and muons. These are identified using standard detector elements such as tracking chambers, electromagnetic and hadronic calorimeters, and muon chambers (details can be found in the references). All experiments remove electron candidates from pair conversion.

The SLD and LEP experiments have data acquisition systems which can record virtually all of the relatively low-rate hadronic $Z$ events. CDF, however, operated with an interaction rate of $\sim$ 300 kHz and simply could not record every event; only those passing a multistage trigger were selected for analysis. This trigger could identify those $b$ events that contained a semileptonic decay by detecting the lepton. However, the fake background at low momentum is very large, and a fairly high lepton $p_t$ threshold of 7.5 GeV was required so that the trigger would

not saturate the data acquisition bandwidth. Dedicated dilepton triggers, either $\mu\mu$ or $e\mu$ were also implemented. For dileptons, the momentum cuts were set much lower (2.0 GeV for the muon and 5.0 GeV for the electron), since requiring two leptons significantly reduces the fake background. The data sets collected with these triggers are referred to as high-$p_t$ single-lepton and low-$p_t$ dilepton, respectively.

In addition to leptons, fragmentation kaons and pions are used by many experiments to tag the initial $b$-quark charge, and kaons from $B$ decay are used by SLD to tag the final charge (see Section 3). ALEPH, CDF and OPAL identify particles based on ionization (dE/dx) deposits in their tracking chambers, while DELPHI and SLD use their ring-imaging Cerenkov detectors.

## 4.2  Dilepton Analyses

### 4.2.1  EVENT SELECTION

The dilepton analyses use a lepton tag as the decay-point tag, as described in Section 3.1, and an opposite-side lepton tag as the production-flavor tag.

Like all opposite-side-type analyses, the dilepton analyses divide candidate events into two hemispheres using one of the methods discussed in Section 3. Events are selected which contain two leptons ($l = e, \mu$) in opposite hemispheres and have a secondary vertex associated with at least one of the leptons. The details of the vertexing algorithms are given below. The leptons are required to pass the cuts on their total momentum (or momentum transverse to the beamline, for CDF) listed in Table 7, which also shows the number of events and number of secondary vertices found in each analysis. In addition, most analyses reject leptons with a $p_t^{\text{rel}}$ below the threshold given, where $p_t^{\text{rel}}$ is the momentum of the lepton transverse to the direction of the associated $b$ jet (with the lepton removed). OPAL rejects leptons based on the output of a neural net, described below.

CDF has three dilepton analyses – two based on the data recorded with the $\mu\mu$ and $e\mu$ triggers and the third based on events passing the single-lepton trigger that have a secondary vertex associated with the lepton. From this sample, events with another so-called soft lepton opposite the trigger lepton are accepted (the algorithm for selecting the lepton is based on the soft lepton tag, or SLT, used in the top quark search).

### 4.2.2  SECONDARY VERTEX RECONSTRUCTION

Because the lepton flavor-tagging method identifies only one possible track from the $B$ decay, all experiments use an inclusive vertex-finding technique to locate the $B$ decay point.

ALEPH [76] approximates that all of the $B$-decay products except the lepton come from the tertiary charm vertex (commonly loosely referred to as the $D$ vertex). They form a grid of possible charm vertex positions. If a track's impact parameter (excluding the lepton) is within $3\sigma$ of a candidate vertex, it is assigned to the vertex. ALEPH calculates the $\chi^2$ difference between assigning all tracks to the primary vertex and allowing some to come from the charm vertex, and the point which maximizes this difference is chosen as the charm vertex. A charm pseudotrack is formed, which passes through this vertex, and its direction is given by the sum of the momentum of the charged particles in the vertex.

Table 7: Summary of selection requirements on the momentum and $p_t^{\mathrm{rel}}$ of the leptons for the dilepton analyses, along with the number of events and reconstructed vertices for each experiment. $< P_r >$ is the mean right-tag probability for the lepton opposite the vertex hemisphere.

| Experiment | Lepton requirements (GeV) | Events/Vertices | $< P_r >$ |
|---|---|---|---|
| ALEPH [76] | $p > 2.0(e), 3.0(\mu); p_t^{\mathrm{rel}} > 1.25$ | 5957/9710 | 0.83 |
| CDF($\mu\mu$) [77] | $p_t > 3.0, p_t^{\mathrm{rel}} > 1.3$ | 5968/5968 | 0.74 |
| CDF($e\mu$) [79] | $E_t > 5.0(e), p_t > 2.5(\mu), p_t^{\mathrm{rel}} > 1.25(\mu)$ | 10180/11844 | 0.74 |
| CDF(SLT) [78] | $p > 7.5(\mathrm{trig}), 2.0(\mathrm{tag}), p_t^{\mathrm{rel}}(\mathrm{weight})$ | 12700/12700 | 0.73 |
| DELPHI [80] | $p > 3.0, p_t^{\mathrm{rel}} > 1.2$ | 4778/4778 | 0.89 |
| L3 [81] | $p > 3.0(e), 4.0(\mu), p_t^{\mathrm{rel}} > 1.0$ | 1490/1928 | 0.84 |
| L3(IP) [81] | same | 2596/− | N/A |
| OPAL [82] | $p > 2.0, \mathrm{neural\ net}$ | 5357/8544 | 0.85 |

This pseudotrack is then vertexed with the lepton to give the position of the secondary, $B$ vertex, which is finally projected onto the $b$-jet direction to give the decay length $L$. This technique allows a $B$ decay length to be determined for every event and is appropriate for samples with high initial $b$ purity, which do not need to rely upon the presence of a well-separated secondary vertex to increase the $b$ purity.

On the other hand, at CDF, the presence of a secondary vertex is a powerful means of rejecting large non-$b$ backgrounds. CDF uses two techniques, both of which involve the projection of the tracks onto the $r\phi$ plane. The dimuon analysis [77] exploits a correlation between the impact parameter, $d$, of a track, and its azimuthal ($\phi$) angle. Tracks coming from a secondary vertex form a line in the $d\phi$ plane, whereas tracks from the primary vertex have both small $d$ and no $d\phi$ correlation. A cluster of tracks from a secondary vertex is formed if at least three tracks with significant separation from the primary vertex ($d/\sigma_d > 2$) forming a line in the $d\phi$ plane can be found. All tracks in this cluster except the muon are vertexed to form a presumed charm vertex, and a pseudotrack is formed, as above, from the summed momentum of these tracks. This is intersected with the muon track to give the $b$ decay vertex position, and hence the flight distance projected onto the $r\phi$ plane, $L_{xy}$.

The second method is used for the high-$p_t$-SLT [78] and $e\mu$ analyses [79]. First, candidate tracks from the $b$ jet are selected if they have significant impact parameters to the primary vertex (this is not required for the lepton track which is presumed to come from the $b$ decay). Combinations of at least three tracks consistent with coming from the same point are candidate vertices. If no such vertex is found, stricter impact parameter and $p_t$ cuts are imposed. All tracks are combined to a common vertex and tracks which contribute too much to the vertex $\chi^2$ are removed. The process is repeated until a good fit is obtained. This technique does not attempt to separate a tertiary charm vertex from the $b$ vertex, but instead combines all tracks into one. The efficiency of these algorithms necessarily falls off steeply as the secondary vertex approaches the primary vertex.

DELPHI [80] uses yet another method, effective on high $b$ purity samples, which attempts to identify tracks from the charm vertex by mass rather than position. Excluding the lepton, they cluster particles in the jet associated with

the lepton using a jet clustering algorithm optimized for this purpose (all tracks are given the pion mass). Within a cluster, particles are ordered by decreasing values of their pseudorapidity relative to the cluster direction. Tracks with the largest pseudorapidity and $p > 500$MeV are kept until the mass of the resulting system exceeds 2.2 GeV, and any such system with a large angle ($> 500$ mrad relative to the $b$ jet) is discarded. The remaining tracks in the cluster are vertexed (in the $r\phi$ plane), and the charm pseudotrack is formed and intersected with the lepton to give the $r\phi$ projection of the $b$ decay point. The decay point with the largest significance of separation from the primary vertex is chosen. The procedure is then repeated once, using the charm vertex tracks as the cluster seed and including neutral particles in the $b$ jet, to obtain the $b$ decay length, $L_{xy}$.

L3 [81] uses a track-based approach similar to CDF. They search for a secondary vertex using the lepton and tracks that are not consistent with coming from the primary vertex. An acceptable vertex can be formed in approximately 70% of the events, with a loss of efficiency when the secondary vertex is near the primary. L3 also has a unique analysis that does not attempt to find a secondary vertex at all; rather it looks at the charge correlation of the two leptons as a function of the impact parameters to the beam axis (IP) of the leptons.

OPAL [82] uses a technique in which tracks, ordered by decreasing significance of separation from the primary, are intersected (in the $r\phi$ plane) with the lepton track to form secondary vertex seeds. Additional tracks that are more consistent with the secondary seed than the primary vertex are added, and a candidate secondary vertex is chosen based on its position and the number of associated tracks. Quality requirements based on the error on the vertex position, the invariant mass of the tracks in the vertex, etc., reduce the efficiency to approximately 70%, though in a decay-length-independent manner.

### 4.2.3  *B* MOMENTUM DETERMINATION

ALEPH estimates the $B$ meson momentum as follows. First, the (vector) sum of the momentum of the lepton and charm vertex tracks is formed. Second, a fraction of the neutral energy in the hemisphere, determined from Monte Carlo studies to be 0.68, is added to this momentum. Finally, the momentum of the neutrino is approximately accounted for by adding the missing momentum in the lepton hemisphere, estimated as the difference between the sum of all visible energy in the hemisphere and the beam momentum.

CDF begins with the total $p_t^{cl}$ in the vertex cluster, and scales it to $p_t^{cl}/\mathcal{K}$, where $\mathcal{K}$ is a factor that is determined from Monte Carlo and parameterized as a function of the observed $p_t^{cl}$ and the mass of the tracks in the cluster.

DELPHI starts with an estimate of the $B$ momentum given by the total energy-momentum reconstructed in its hemisphere minus that of all particles not included as part of the $B$ in the vertex algorithm, which already contains contributions from neutral energy. A single neutrino's energy and direction is then solved for by assuming it is the only missing particle in the event. The visible energy and momentum are rescaled by a factor $\alpha = 1.13$, which gives the optimal resolution. The energy $E_\nu$ is added to the $B$ momentum if the direction of the neutrino momentum is within 400 mrad of the $l$-charm direction previously determined. Finally this estimate is adjusted to to give the correct mean value, based on the Monte Carlo prediction.

L3 does not estimate the $B$ momentum event-by-event but rather uses a constant value of $0.85E_{\text{beam}}$.

OPAL uses a very different technique from the other experiments. OPAL first calculates the invariant mass of all the objects in the event that are not in the $b$ jet, assuming all charged tracks are pions. This mass is corrected slightly based on the inverse of the total reconstructed energy (which should be $M_Z$) to improve the resolution. The energy of the $b$ jet, including the neutrino, is determined, assuming the $Z^0$ undergoes a two-body decay to particles of mass 5.3 GeV. Some of this energy is removed as follows. For charged tracks, a weight is calculated based on: the probability that the track comes from the secondary, rather than the primary, vertex; the track's momentum, and the angle of the track to the estimated charm-hadron direction. For neutral clusters, the weight is based on the angle between the cluster direction and the charm direction alone. The weighted sum of charged and neutral energies in the jet is then subtracted from the full $b$ jet energy to give the estimated $B$ meson momentum. This technique results in an excellent fractional resolution of approximately 12%, with the core being closer to 7%. It was found that the resolution in dilepton events, with two missing neutrinos, was not degraded significantly.

### 4.2.4   FIT METHOD AND RESULTS
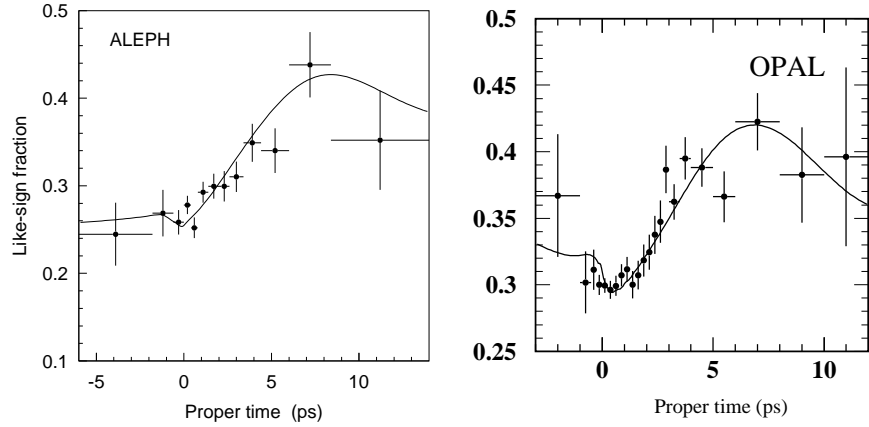


Figure 9: Fraction of like-sign events vs. proper time. *Left:* dilepton events from ALEPH *Right:* lepton-jet charge events (selected with $|Q_{\text{tag}}| > 2$) from OPAL. The curves represent fits to the data.

The analysis variable for dilepton events is the product of the lepton signs $q_c(t) = q_l^{\text{opp}} \cdot q_l^{\text{vtx}}(t)$. Opposite-sign leptons make up the "unmixed" $DP+$ sample, and same-sign leptons the $DP-$ sample. The proper time dependence of these samples is predicted using the resolution-smeared, efficiency-weighted distributions of the sample components, weighted by their relative abundance, combined with the probability of getting the correct charge assignments (as described in Section 3.3). If secondary vertices are found for both leptons, the joint probability of finding decays at their respective decay times is constructed. The CDF $\mu\mu$ and $e\mu$ analyses use these predictions in a binned least-squares fit, while the rest of the analyses use unbinned likelihood fits.

The $b$ purity of these samples is fairly high, ranging from 94% for DELPHI,

to 98% for ALEPH, with the CDF samples typically at 80%. The remaining events are either $c\overline{c}$, or fakes. $c\overline{c}$ events represent between 0.5% of the sample (for ALEPH, L3 and the low-$p_t$ CDF samples) and 8% (for CDF). The fake background, including fake leptons from both $b\overline{b}$ and non-$b$ sources, constitutes a few percent – up to 15% at CDF. Approximately 38% of the leptons in $b\overline{b}$ events are estimated to come from $B^0$ mesons.

Many of the following analyses use the opposite-side lepton as a production flavor tag, so a few words concerning its effectiveness are in order. The lepton opposite the vertex side can mistag the $b$ in that hemisphere if the the $b$ hadronizes into a $B^0$ or $B_s^0$ and mixes, if the lepton comes from a cascade $b \to c \to l$ decay, or if the lepton is a fake. Denoting the fractions of the various $B$ hadrons by their light quarks as $f_d, f_u, f_s$, and $f_{\Lambda_b}$ for all the baryons, the contribution to the mistag rate from mixed $B^0$ and $B_s^0$ mesons is $\overline{\chi} = f_d \chi_d + f_s \chi_s$ (ingnoring for the moment those events with vertices on both sides). Since $\chi_d = 0.172 \pm 0.010$ [6], $\chi_s \approx 0.5$ (see Section 5), then $\overline{\chi} = 0.118 \pm 0.006$ [6], with contributions from $f_d \approx 40\%$ and $f_s \approx 11\%$. Thus, even though $B_s^0$ mesons only represent 11% of the $B$ mesons produced, they contribute substantially to the overall mistag rate due to their fast oscillation rate.

The other significant source of mistags are the wrong-sign cascade decays. Table 5 lists the sources of leptons which do not come directly from the $b$ quark. Among these, the first two make up the bulk of the events. The right-sign probability, then, is $P_r = 1 - \overline{\chi} - f_{bc}^{ws}$, where $f_{bc}^{ws}$ is the fraction of wrong-sign cascade decays from the mix of $b$ hadron types in the hemisphere. In the absence of any cuts on the lepton, this fraction is comparable to the contribution from direct $b$ decays. To reduce this fraction, all experiments try to remove, or assign small weights to, events with a large probability of having a lepton from a cascade decay. Each experiment uses the momentum of the lepton referenced to the direction of the jet (with the lepton removed), $p_t^{rel}$. Leptons from direct $b$ decay can be produced with significant mometum perpendicular to the $b$ flight direction due to the large $b$-quark mass, whereas those from the charm decay typically have less transverse momentum. For the same reason, cascade leptons also tend to be less well isolated from the core of the jet, and because they come from further down the decay chain, their mean momentum is lower. OPAL feeds these three variables into a neural net analysis, which is estimated to be 40% more effective at rejecting cascade decays than simple cuts on $p$ and $p_t^{rel}$. These cuts reduce the right-sign cascade component, $b \to \overline{c} \to l$, even more, since these events are of the $B \to D\overline{D}$ type, and the momentum available for the lepton is even smaller than in other cascade decays. These end up accounting for $12 - 25\%$ of the cascade decays.

The SLT analysis of CDF, rather than simply cutting on $p_t^{rel}$, weights events according to the probability that the flavor tag lepton identifies the $b$ charge correctly (parameterized as a function of $p_t^{rel}$). The other analyses use the mean probability for each event.

The sample composition on the vertex side need not be the same as on the opposite side. For example, in the CDF SLT analysis the vertex-side lepton has $p_t > 7.5$ GeV, while the opposite-side lepton satisfies $p_t > 2.0$ GeV. In addition, vertexing algorithms with efficiencies that decrease near the primary vertex increase the fraction of the longer-lived sample components.

Figure 9 shows the like-sign fraction vs. proper time for the ALEPH analysis. Table 8 lists the results of fits to the oscillation frequency $\Delta m_d$, along with the

source and fractional variation of the parameters which contribute most to the systematic error. The $B_s^0$ and cascade fractions are important since they dominate the mistag rate. The ratio of $B^+$ to $B^0$ lifetime is also important, since this can enhance or suppress the fraction of $B^+$ to $B^0$ at long lifetimes, altering the shape of the like-sign fraction and distorting the $\Delta m$ measurement.

Table 8: Fit results and dominant systematic error for dilepton analyses.

| Experiment | Dominant Systematic | $\delta(\Delta m_d)$ (ps$^{-1}$) | $\Delta m_d$ (ps$^{-1}$) |
|---|---|---|---|
| ALEPH | cascade frac. (15%) <br> $B_s^0$ frac. (9%) | $\pm 0.025$ <br> $\pm 0.019$ | $0.452 \pm 0.039 \pm 0.044$ |
| CDF($\mu\mu$) | cascade frac. (15%) <br> $\tau_b$, sample comp. | $\pm 0.048$ <br> $\pm 0.043$ | $0.503 \pm 0.064 \pm 0.071$ |
| CDF($e\mu$) | cascade frac. (15%) <br> $b$-Baryon fract. (40%) | $\pm 0.030$ <br> $\pm 0.021$ | $0.450 \pm 0.045 \pm 0.051$ |
| CDF(SLT) | cascade frac. (25%) <br> $\tau_{B^+}/\tau_{B^0}$(5%) <br> charm dilut. (100%) | $\pm 0.004$ <br> $\pm 0.021$ <br> $\pm 0.032$ | $0.500 \pm 0.052 \pm 0.043$ <br> (includes Jet Q tags) |
| DELPHI | cascade frac. (6%) <br> $\tau_{B^+}/\tau_{B^0}$(4%) | $\pm 0.032$ <br> $\pm 0.017$ | $0.480 \pm 0.040 \pm 0.051$ |
| L3 | cascade frac. (15%) <br> $B_s^0$ frac. (14%) | $\pm 0.022$ <br> $\pm 0.015$ | $0.458 \pm 0.046 \pm 0.032$ |
| L3(IP) | cascade frac. (15%) <br> $B_s^0$ frac. (14%) | $\pm 0.037$ <br> $\pm 0.026$ | $0.472 \pm 0.049 \pm 0.053$ |
| OPAL | cascade frac. (15%) <br> $\tau_{B^+}/\tau_{B^0}$(6%) | $\pm 0.011$ <br> $\pm 0.023$ | $0.430 \pm 0.043 \pm 0.030$ |

## 4.3   Lepton-Jet-Charge Analyses

This section introduces the second commonly used opposite-side production flavor tag, the jet charge, as it is used in events with a lepton-tagged decay vertex.

### 4.3.1   EVENT SELECTION AND PROPER TIME MEASUREMENT

Events with a single lepton passing each experiment's standard lepton selection are considered. The lepton identification procedure follows that of the dilepton analyses for ALEPH [76], DELPHI [80], L3 [81] and OPAL [83], except that ALEPH raises the momentum cut on the electron to 3 GeV to match that of the muon. The CDF [78] data sample is the same as that of the lepton-SLT analysis described in the previous section. These experiments search for a secondary vertex in the lepton hemisphere and convert to proper time, as described above for the dilepton analyses.

SLD select leptons with $p > 2.0$ GeV and $p_t^{\rm rel} > 0.4$ GeV ($p_t^{\rm rel} > 0.8$ GeV) for the inclusive lepton [84] and lepton-$D$ [85] samples, defined below.

SLD's vertexing algorithm involves combining all the (well-measured) tracks in a hemisphere. The position of the $b$ vertex relative to the primary is estimated by $\vec{L} = \frac{\sum W_i \vec{X}_i}{\sum W_i}$, where $\vec{X}_i$ is the vector from the primary vertex to the point on the lepton track that is closest to the track of particle $i$, and $W_i$ is a weighting function that is the product of three weights. The first function suppresses tracks which are consistent with the primary vertex; it depends upon the impact parameter (three-dimensional) of the track. The function approaches zero for small impact parameters and becomes roughly constant for large impact parameters. The second function suppresses tracks that have a poorly measured intersection with the lepton, and the third gives more significance to tracks for which this intersection is close to the intersection of the lepton track and jet direction. This method results in a vertex resolution with a core distribution of width $170\mu$m.

SLD uses a second method in which a $D$-like vertex is searched for inclusively in the lepton hemisphere. Tracks are first classified as non-primary if they have an impact parameter (three-dimensional) to the primary vertex of $> 3.5\sigma$ and a momentum greater than 0.8 GeV. The non-primary tracks (excluding the lepton, which is presumed to come from the $B$ decay) are vertexed, and the resultant pseudotrack defined by the $D$ vertex and momentum is intersected with the lepton. If this system forms a valid vertex, an attempt is made to add a primary track first to the $B$ then $D$ vertex. If this whole procedure fails, SLD searches for a single non-primary track that forms a valid vertex with the lepton. If this succeeds, the remaining tracks are vertexed to form the $D$ as above. Although this is not an exclusive $D$ reconstruction, we denote this analysis by $lD$ to distinguish it from the above method.

Apart from the typical vertex quality requirements, SLD also requires that the total charge of the tracks in both the $D$ and $B$ vertices be -1, 0 or 1. Because of the high probability of correctly assigning tracks to secondary and tertiary vertices from the excellent tracking detectors, the SLD group can use the total charge at the $B$ vertex to select $B^0$ events with a high efficiency, while excluding a large fraction of the $B^+$ events. They estimate the sample contains 60.8% $B^0$ and only 19.6% $B^+$.

SLD obtains the $B$ momentum for the $l$ hemisphere as follows. An attempt is made to include charged tracks only from the $B$ decay by ordering the tracks in the jet based on the significance of their impact parameter to the primary vertex (in three dimensions). These tracks are combined until the invariant mass of the combination exceeds 2.0 GeV. This set of tracks, along with the lepton, give the charged particle contribution to the $B$ momentum. The contribution from the neutrino is given by the difference between the beam energy and the total visible energy in the lepton hemisphere. Neutral energy is added based on the parameterization $E_{B^0} = 0.7E_0 + 0.01E_0^2$, where $E_{B^0}$ is the neutral energy to be assigned to the $B$ meson, and $E_0$ is the total visible neutral energy in the $b$ jet. This method results in a core resolution of 8% on the $B$ momentum.

The SLD $lD$-jet charge analysis fits to the decay length distribution directly, and so does not use event-by-event momentum information.

The event selection is summarized in Table 9.

Table 9: Jet charge tag definitions for the inclusive lepton analyses, along with the right-sign probability parameterization as a function of $Q_{\text{tag}}$. The mean value of $P_r$ is quoted for rough comparison.

| Experiment | $Q_{\text{tag}}$ Tag Defn. | $P_r(Q_{\text{tag}})$ (or Event Weight) | Events | $<P_r>$ |
|---|---|---|---|---|
| ALEPH [76] | $Q_H^{\text{opp}}(0.5)$ | $w = \lvert Q_{\text{tag}} \rvert$ $(P_r = \frac{1+\lvert Q_{\text{tag}} \rvert}{2})$ | 62,320 | 0.53 |
| CDF [78] | $Q_J^{\text{opp}}(1)$ | $\frac{1+N_D \lvert Q_{\text{tag}} \rvert D_{\max}}{2}$ | 120,700 | 0.56 |
| DELPHI [80] | $Q_H^{\text{opp}}(0.6)$ | $<P_r>, \quad (\lvert Q_{\text{tag}} \rvert > 0.1)$ | 60,381 | 0.69 |
| L3 [81] | $Q_H^{\text{opp}}(0.4) - Q_H^{\text{vtx}}(0)$ | $<P_r>, \quad (\lvert Q_{\text{tag}} \rvert > 0.12)$ | 8,707 | 0.72 |
| OPAL [83] | $Q_J^{\text{vtx}}(0) - 10 Q_J^{\text{opp}}(1)$ | From MC | 94,843 | |
| SLD [84] | $Q_H^{\text{opp}}(0.5)$ | $\frac{1}{1 + e^{-0.26 Q_{\text{tag}}}}$ | 2,609 | 0.68 |
| SLD [85](lD) | $Q_H^{\text{opp}}(0.5)$ | $\frac{1}{1 + e^{-0.32 Q_{\text{tag}}}}$ | 584 | 0.68 |

## 4.3.2   JET-CHARGE FLAVOR TAG

The basic building block for the jet charge tag $Q_{\text{tag}}$ is either the weighted hemisphere or $b$-jet charge, commonly defined as

$$Q_{H,J}(\kappa) \;=\; \frac{\sum_i q_i p_{l,i}^{\kappa}}{\sum_i p_{l,i}^{\kappa}}, \qquad p_{l,i} = \vec{p}_i \cdot \hat{a}$$

Here $\hat{a}$ is the unit vector that is used to divide the event into two hemispheres (thrust, sphericity or $b$-jet axis), $p_{l,i}$ is the $i^{\text{th}}$ track's momentum component along this axis, and the sum runs over all charged tracks in the hemisphere ($Q_H$, used by ALEPH, DELPHI, L3 and SLD) or jet ($Q_J$ used by CDF, OPAL). OPAL normalizes to the beam energy $E_{\text{beam}}^{\kappa}$ rather than $\sum_i p_{l,i}$; SLD has no normalization in the denominator; and L3 uses $C(\phi) q_i$ in place of $q_i$, where $C(\phi)$ is a function that deweights tracks that pass near an anode wire of the tracking chamber, for which the charge is not well determined. $\kappa$ is chosen to optimize the power of $Q_H$ to distinguish $b$ from $\bar{b}$ quarks. For all experiments except SLD, $-1 < Q_H < 1$. $Q_H(\kappa = 0)$ is simply the mean track charge in the hemisphere, $\sum_{i=1}^{n} q_i/n$ (sum of charges for SLD), and so is typically small. $Q_H(\kappa = \infty)$ is simply the charge of the track with the largest longitudinal momentum fraction, so $Q_H(\infty) = \pm 1$.

  $Q_H$ is the basic building block for the jet-charge $Q_{\text{tag}}$; however several experiments use a linear combination of the two hemisphere (jet) charges (with different $\kappa$ weights) as the tag variable. While jet charge is basically an opposite-side tag, the fragmentation tracks on the (possibly mixed) $B$-vertex side still contain information about the $b$-quark flavor at its production, although the correlation is weaker than on the opposite side. Since one is searching for mixing on this side, it is desirable to find a combination that only depends weakly on whether the $B$ meson has mixed (otherwise the right-tag probability would depend upon the mixed/unmixed state of the $B$, complicating the fit). The sum of the charges in the vertex hemisphere, (i.e. $Q_H^{\text{vtx}}(\kappa = 0)$) is relatively insensitive to whether the $B$ meson mixed, as only neutral mesons can mix. L3, for example, finds that the combination $Q_{\text{tag}} = Q_H^{\text{opp}}(\kappa = 0.4) - Q_H^{\text{vtx}}(\kappa = 0)$ has maximal analyzing

power. The combinations used by other experiments in the $l$-$Q_{\text{tag}}$ analyses are given in Table 9. SLD combines the jet-charge tag with their polarization tag described in Section 3.2. Finally, note that for OPAL, $Q_{\text{tag}}$ has been defined with the opposite sign and a scale factor of approximately ten. Figure 10$b$ shows the measured jet-charge distribution from SLD, along with the underlying $b$ and $\overline{b}$ predictions from Monte Carlo studies.
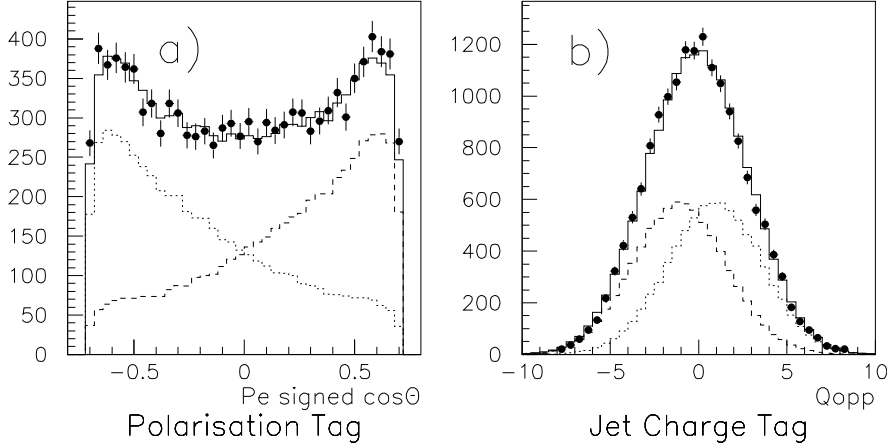


Figure 10: $b$-quark production flavor tags for the SLD experiment. *Left:* Polarization tag, based on the polar angle of the event axis. *Right:* Jet charge. The dashed lines are the distributions for $b$ and $\overline{b}$ quarks, and the points are the data.

One of the advantages of the jet-charge tag is its high efficiency. At LEP and SLD, the $b$ jets are mostly produced back-to-back. Since the vertex-side jet must have tracks passing through the silicon detectors, it is well within the solid-angle covered by each experiment, and the opposite-side $b$ jet is similarly well contained. A comparison between the $l$-$l$ and $l$-$Q_{\text{tag}}$ samples shows the latter contain approximately ten times more events.

At CDF, however, the initial $b\overline{b}$ pair is not produced at rest, but with a potentially large momentum along the beam axis. A typical rapidity separation between the two $b$ jets is $\Delta\eta \approx 1$. The tracking chambers cover the region $-1 < \eta < 1$, so quite often the vertexed $b$ jet is contained within the acceptance of the tracking system while the other $b$ jet is not. Even worse, another, non-$b$ jet may be present, diluting the tag's effectiveness. These effects result in the average right-tag probability of 56%, rather than $\sim$ 67% that is typical for the LEP and SLD experiments. However, a search for a secondary vertex in the non-lepton jet is performed. The subset of events with such a vertex has a greatly enhanced right-tag probability, namely 63%.

### 4.3.3   FIT METHOD AND RESULTS

The sign of the tag variable $Q_{\text{tag}}$ is used to categorize the sign of the $b$-quark charge in the opposite hemisphere. The analyses are based on the time dependence of $q_c(t) = \text{sign}(Q_{\text{tag}}) \cdot \text{sign}(q_l(t))$. CDF, DELPHI, L3, OPAL and SLD perform likelihood fits to the same- and opposite-sign data, while ALEPH per-

forms a binned-$\chi^2$ fit. The quantity $< -Q_H \cdot q_l(t) >$ is formed in 0.2 ps bins of proper time. This is simply the standard charge correlation function $q_c(t)$ with each event weighted by $|Q_H|$. Although weighting by the dilution $\mathcal{D} = 2P_r - 1$ is optimal, note that CDF finds that for their procedure, in fact, $\mathcal{D} \sim Q_H$.

Events with $|Q_{\text{tag}}|$ near zero clearly have little power to separate opposite hemisphere $b$ from $\bar{b}$ quarks. DELPHI and L3 simply cut out events with $|Q_{\text{tag}}| < 0.1$ and $|Q_{\text{tag}}| < 0.12$ respectively.

OPAL and SLD, on the other hand, parameterize $P_r(Q_{\text{tag}})$ based on Monte Carlo studies. For example, the separation between $Q_{\text{tag}}$ for $b$ and $\bar{b}$ jets at SLD is shown in Figure 10, which can be combined to give the $b$ probability as a function of measured $Q_{\text{tag}}$. CDF prefers to measure the shape of $P_r(Q_{\text{tag}})$ from data. They measure the raw dilution, $\mathcal{D}_{\text{raw}} = (N_{SS} - N_{OS})/(N_{SS} + N_{OS})$, where SS refers to same-sign and OS refers to opposite-sign lepton and jet-charge. They find that the raw dilution can be parameterized by a simple straight line, $D_{\text{raw}}(|Q_{\text{tag}}|) = \alpha|Q_{\text{tag}}|$. Since the raw dilution is lower than the true dilution due to mixing and cascade semileptonic decays, the final fit includes an overall normalization, $\mathcal{D} = N_D D_{raw}(|Q_{\text{tag}}|)$. The normalization is $N_D \approx 1.5 - 2.0$, depending on the subsample.

In addition to estimating the production tag right-sign probability for each event, OPAL also estimates the sample composition fractions event-by-event based on the value of the output of the neural net that is used to select lepton candidates, as described in Section 4.2.

The quality of the vertex-side lepton tag depends on the parameters discussed in the dilepton section. The sample composition on the lepton side is very similar to the dilepton case. The right-sign probability for this decay-point tag is typically estimated from the Monte Carlo predictions for the $B_s^0$ and cascade fractions. The tag purity of the jet charge is extracted from the data (see Table 9 for a summary and the number of events in each analysis). The systematic errors are dominated by the same effects which distort the relative amount of mixed and unmixed events as a function of the proper decay time, namely the sample composition parameters and the $B^+$-$B^0$ lifetime ratio. The results of the fits are summarized in Figure 12 and Table 10.

### 4.4  $D^* - l$ and $D^*-$Jet Charge Analyses

As described in Section 3.1, a reconstructed $D^{*+}$ meson can be used as a very pure tag of the charge of the $b$ quark at its decay point; moreover, events with $D^{*+}$ are also enriched in $B^0$ content. Further, the tracks forming the $D^{*+}$ can be vertexed and used to estimate the $B$ decay length.

The experiments search for the following decay modes: $D^{*+} \to D^0\pi_*^+$, with $D^0 \to K^-\pi^+, K^-\pi^+\pi^0$, or $K^-\pi^+\pi^+\pi^-$. All experiments exploit the small Q-value of the $D^{*+}$ decay: $M_{D^{*+}} - M_{D^0} - M_{\pi^+} = 5.8$ MeV, meaning the decay products are basically at rest in the $D^{*+}$ rest frame, which greatly suppresses the combinatorial background. The quantity $\Delta M = M(\text{all found decay products including } \pi_*^+) - M(\text{all found decay products excluding } \pi_*^+)$ is very insensitive to the momentum resolution of an individual track, since resolution effects mostly cancel in the difference. This is true even to the extent that the $\pi^0$ can be missing and a clean $D^*$ signal extracted.

ALEPH [76] fully reconstructs the three $D^0$ modes listed, including the $\pi^0 \to \gamma\gamma$. CDF [86] reconstructs the $D^{*+}$ only in the $D^0 \to K^-\pi^+$ mode, but adds in

Table 10: Fit results and dominant systematic errors for the lepton-jet charge analyses.

| Experiment | Dominant Systematic | $\delta(\Delta m_d)$ (ps$^{-1}$) | $\Delta m_d$ (ps$^{-1}$) |
|---|---|---|---|
| ALEPH | $\tau_{B^0}(3\%)$ $\tau_{B^+}(3\%)$ | $\pm 0.012$ $\pm 0.016$ | $0.404 \pm 0.045 \pm 0.027$ |
| CDF | $\tau_{B^+}/\tau_{B^0}(5\%)$ | $0.021$ | $0.500 \pm 0.052 \pm 0.043$ (includes $l$ tags) |
| DELPHI | right-sign frac., fakes (2%) fake frac. (20%) $b$-Baryon frac. (33%) | $\pm 0.014$ $\pm 0.010$ $\pm 0.010$ | $0.493 \pm 0.042 \pm 0.027$ |
| L3 | $<P_r> (2\%)$ | $\pm 0.037$ | $0.437 \pm 0.043 \pm 0.044$ |
| OPAL | $B_s^0$ fraction (17%) $\tau_{\Lambda_b}/\tau_{B^0}(8\%)$ | $\pm 0.011$ $\pm 0.007$ | $0.444 \pm 0.029 \pm 0.020$ |
| SLD($l$) | $B_s^0$ frac. (26%) $b$-Baryon frac. (42%) | $\pm 0.014$ $\pm 0.015$ | $0.520 \pm 0.072 \pm 0.035$ |
| SLD($lD$) | Fit binning and range Tracking efficiency | $\pm 0.032$ $\pm 0.017$ | $0.452 \pm 0.074 \pm 0.049$ |

the $D^+ \to K^- \pi^+ \pi^+$ decay. DELPHI [80] does not reconstruct the $\pi^0$, resulting in larger background. OPAL [87] finds the signal/background for the $K\pi(\pi^0)$ and $K3\pi$ modes to be too high, and excludes them from the analysis. Table 11 lists the modes reconstructed, as well as the signal-to-background ratio for each analysis.

### 4.4.1 FLAVOR TAGGING

The vertex-side $b$ quark flavor at its decay point has the opposite sign of the reconstructed $D^{*\pm}$. The production flavor tags are either the opposite-side lepton, or opposite-side jet-charge tags which were introduced above.

### 4.4.2 DECAY LENGTH AND PROPER TIME DETERMINATION

These analyses have the advantage of a built-in decay vertex, namely that given by vertexing the $D^0$ decay products. Unfortunately, intersecting this with the $\pi_*^+$ from the $D^{*+}$ decay to obtain the $B$ decay point is not viable; the pion is emitted almost at rest in the $D^{*+}$ frame, and so its direction is the same as the $D^0$ direction in the laboratory frame. All analyses begin by vertexing the $D$ decay tracks.

Only OPAL attempts to explicitly find the $B$ decay vertex, as follows. After vertexing the tracks forming the $D^{*+}$ and creating a pseudotrack passing through this vertex with the $D^{*+}$ momentum, OPAL attempts to combine this pseudotrack with other tracks in the hemisphere that are likely to have come from the $B$ decay. These tracks are selected based on their momentum and angle relative to the $D^{*+}$ pseudotrack direction, and they must have an intersection point with

Table 11: Overview of the decay modes, sample sizes, flavor tag and mean right-sign tag probability for the $D^*$-$l$ and $D^*$-jet charge analyses.

| Experiment | $D^0/D^+$ Modes | Sig/Bgnd | Prod. Tag | $<P_r>$ |
|---|---|---|---|---|
| ALEPH [76] | $K^-\pi^+$ $K^-\pi^+\pi^0$ $K^-\pi^+\pi^+\pi^-$ | 840/565 | $l$ $(p > 3.0$ GeV $p_t^{\mathrm{rel}} > 0.75$ GeV) | 0.79 |
| ALEPH [76] | $K^-\pi^+$ $K^-\pi^+\pi^0$ $K^-\pi^+\pi^+\pi^-$ | 1555/1096 | Jet Q $Q_H(0.5) - 0.08\sum q_i$ $|Q_{\mathrm{tag}}| > 0.1$ | 0.76 |
| CDF [86] | $K^-\pi^+$ $K^-\pi^+\pi^+$ | 358/520 460/510 | $l$ $p_t > 8.0$ GeV $p_t^{\mathrm{rel}} > 1.5$ GeV | 0.79 |
| DELPHI [80] | $K^-\pi^+$ $K^-\pi^+(\pi^0)$ $K^-\pi^+\pi^+\pi^-$ | 1554/745 1370/1121 1288/3904 | Jet Q$^\ddagger$ | 0.65$^\dagger$ |
| OPAL [87] | $K^-\pi^+$ | 253/95 | $l$ Neural net$^*$ | 0.79 |

$^\dagger$Includes wrong-sign $D^*$ mistags.  $^\ddagger$See Table 9.  $^*$Approximately equivalent to cuts of $p > 3.0$ GeV and $p_t^{\mathrm{rel}} > 0.75$ GeV.

the pseudotrack that is consistent with a $b$ decay. The $B$ momentum is estimated using the same technique as in the OPAL dilepton analysis.

CDF estimates the most likely $B$ decay position based on its distance to the $D$ vertex, the momentum of the $D$, and knowledge of the $D$ and $B$ lifetimes. The $B$ momentum is estimated from the momentum of the $D$ meson, using a correspondence determined from Monte Carlo studies.

ALEPH and DELPHI fit the charge correlation to variables described below.

### 4.4.3   FIT PROCEDURE AND RESULTS

ALEPH fits directly to the charge correlation of the $D^*$ sign and opposite hemisphere sign ($l$ or jet charge) as a function of the $D^0$ decay distance, $q_c(L_{D^0})$. This involves combining both $B$ and $D$ decay time and momentum distributions in Monte Carlo studies to generate the predicted decay length distributions used in the likelihood fit.

DELPHI fits to the charge correlation as a function of the sum of the proper decay times of the $B$ and $D$ meson. For $l_B$=flight distance of the $B$, $l_D$=flight distance of the $D$, and $l \approx l_B + l_D$ = total decay length to the $D$ vertex, one defines a proper time $t = t_B + t_D = \frac{m_B}{p_B}l_B + \frac{m_D}{p_D}l_D \approx \frac{m_B}{p_B}l$. This approximation is good to about a percent. DELPHI takes the average value $p_B = 0.7E_{\mathrm{beam}}$ to convert $l$ to $t$.

CDF fits to the mean charge correlation in eight bins of proper time while OPAL performs an unbinned likelihood fit as described previously.

The vertex-side tag (the sign of the $D^{(*)}$), has a very high purity – the main source of mistags comes from $B \to D\overline{D_s}$ events (see Section 3.1). The $B^0$ fraction

in these samples is high, with a typical value of 83%. None of the experiments measure the production tag's right-sign probability event-by-event; instead they use the mean value $< P_r >$, which is left as a free parameter in the fits. The values of $< P_r >$ are summarized in Table 11.

Figure 12 and Table 12 summarize the results of the fits.

Table 12: Fit results and dominant systematic errors for the $D^*$-$l$ and $D^*$-jet charge analyses.

| Experiment | Dominant Systematic | $\delta(\Delta m_d)$ (ps$^{-1}$) | $\Delta m_d$ (ps$^{-1}$) |
|---|---|---|---|
| ALEPH (combined) | cascade frac. (15%) $B_s^0$ frac. (9%) | ±0.025 ±0.019 | $0.482 \pm 0.044 \pm 0.024$ |
| CDF | $B^+$ frac. (35%) $\tau_c$(25%) | ±0.037 ±0.021 | $0.562 \pm 0.068 \pm 0.050$ |
| DELPHI | cascade frac. (6%) $\tau_{B^+}/\tau_{B^0}$(4%) | ±0.032 ±0.017 | $0.523 \pm 0.072 \pm 0.043$ |
| OPAL | cascade frac. (15%) $\tau_{B^+}/\tau_{B^0}$(6%) | ±0.011 ±0.023 | $0.567 \pm 0.089 \pm 0.029$ |

## 4.5  $D^* l - l$ and $D^* l - Jet$ Charge Analyses

CDF [88], DELPHI [80] and OPAL [87] perform analyses in which the vertex side of the event has both a $D^{*+}$ and a lepton. A semileptonic $B$ decay involving a $D^*$ has the lepton sign opposite the $D^*$ sign, e.g. $D^{*+}l^-$, whereas combinatorial background contributes to both opposite-sign and same-sign. The production flavor tags used are the opposite-side lepton (CDF), or the opposite-side jet charge (DELPHI, OPAL).

### 4.5.1  EVENT SELECTION

CDF searches for $D^{*+}$ candidates in the low-$p_t$ dilepton samples described previously, using the same decay modes described in the above section and similar techniques. OPAL adds the $D^0 \rightarrow K^- \pi^+ (\pi^0)$ decay. Neither experiment reconstructs the $\pi^0$.

DELPHI does not actually reconstruct the $D^{*+}$ at all; they only try to find the $\pi_*^+$ from the decay. Tracks likely to belong to the $B$ decay are selected by the inclusive vertex method in the lepton jet, exactly as described in the dilepton analysis in Section 4.2. Just as in the case of the missing $\pi^0$, forming the mass difference $\Delta M = M$(All found decay products including $\pi_*^+$)$-M$(All found decay products excluding $\pi_*^+$) allows a relatively pure $D^{*+}$ signal to be extracted even if several $B$ decay products are missing. The larger background can be controlled via the extra handle of the charge correlation of the lepton with the $\pi_*$.

Table 13 summarizes the event selection.

### 4.5.2 DECAY LENGTH AND MOMENTUM MEASUREMENTS

In $D^*l$ analyses, the lepton track, which is presumed to come from the $B$ meson, can be used to reconstruct the $B$ decay point. The $D^0$ decay products are vertexed to reconstruct the $D$ decay point, and a pseudotrack is formed with the $D$ momentum passing through this point. This pseudotrack is intersected with the lepton trajectory to give the $B$ decay position. OPAL solves for the decay length $L$ using the position of the primary vertex, this secondary $B$ vertex and a constraint based on the direction of the $lD^{*+}$ system. CDF performs a similar vertexing procedure in the $r\phi$ plane and projects the decay length vector formed by the $B$ and primary vertices onto the $l - D^*$ momentum direction to give $L_{xy}$.

OPAL and CDF estimate the $B$ boost as $\beta\gamma = p^{lD^{*+}}/M_B \times \mathcal{K}(p^{lD^{*+}}, M_{lD^{*+}})$, and $\beta\gamma\sin\theta = p_t^{lD^{*+}}/M_B \times <\mathcal{K}>$ respectively, where the correction factor, $\mathcal{K}$, is determined from Monte Carlo studies. The full distribution of $\mathcal{K}$ is used to derive the momentum smearing needed for the likelihood functions. Typical resolutions on the decay time are 15-20%.

DELPHI converts the decay length found by the inclusive vertex algorithm to proper time exactly as in the dilepton analysis.

Table 13: Overview of the decay modes, sample sizes, flavor tag and mean right-sign tag probability for the $D^*l$-$l$ and $D^*l$-jet charge analyses.

| Experiment | $D^0$ Modes | Sig/Bgnd | Production Tag | $<P_r>$ |
|---|---|---|---|---|
| CDF [107] | $K^-\pi^+$ <br> $K^-\pi^+(\pi^0)$ <br> $K^-\pi^+\pi^+\pi^-$ | 167/49 <br> 190/226 <br> 173/83 | $l^\dagger$ | 0.67 |
| DELPHI [80] | $\pi_*^+ - l$ | 4132/1823 | Jet Q$^\ddagger$ | 0.69 |
| OPAL [83] | $K^-\pi^+$ <br> $K^-\pi^+(\pi^0)$ | 406/49 <br> 794/225 | Jet Q$^\ddagger$ <br> ($|Q_{\text{tag}}| > 1$) | 0.72 |

$^\dagger$ See $e\mu$ and $\mu\mu$ in Table 7.  $^\ddagger$ See Table 9.

### 4.5.3 FIT METHOD AND RESULTS

Each experiment performs a likelihood fit to the charge correlation of the charge of the vertex-side lepton and the opposite side lepton or jet-charge sign, as a function of the proper decay time.

The decay-point tag is of course very pure, having both a $D^*$ and lepton requirement. None of the experiments uses an event-by-event measure of the production tag's right-sign probability; instead, they use the mean value $<P_r>$, which is a free parameter in the fits. To increase this mean probability, OPAL and DELPHI remove low $|Q_{\text{tag}}|$ events. The fitted values of $<P_r>$ are given in Table 13.

Figure 11 shows the $DP-$ (mixed) fraction for the DELPHI $\pi_*l - l$ analysis. Figure 12 and Table 14 summarize the results of the fits. The systematic errors are dominated by the uncertainty in the $B^-$ contamination.
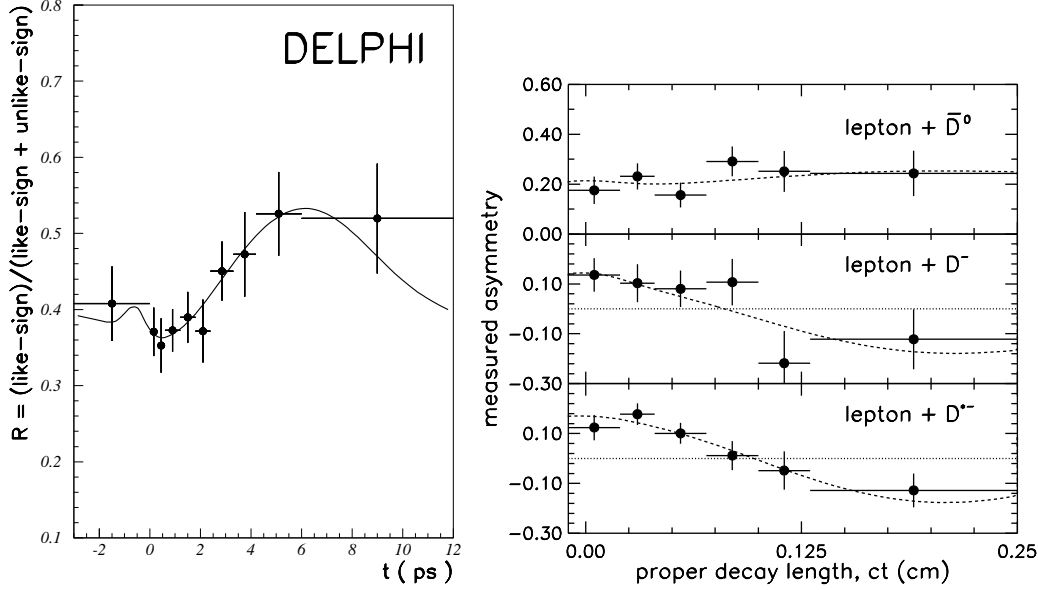
Figure 11: *Left*: Fraction of $\pi_* l - l$ events with like-sign (mixed) leptons, from the DELPHI analysis, and the fit results. *Right*: Right/wrong sign asymmetry from the CDF same-side tagging analysis and fit results.

## *4.6   Other Inclusive Analyses*

There have been several attempts to increase the sample sizes beyond those involving leptons or $D$ tags by using very inclusive methods, both in locating candidate $B$ vertices and in flavor tagging.

Both ALEPH [89] and SLD [90] have selected events using similar inclusive topological vertexing methods. SLD applies kaon and dipole final-state tags, using jet-charge and polarization for the initial tag, while ALEPH uses a highly efficient double jet-charge method.

### 4.6.1   EVENT SELECTION

The vertex algorithms used by ALEPH and SLD, rather than being driven by track intersections, search for viable secondary vertex positions. The ALEPH vertex algorithm is as described in the dilepton section, and is based on the $\chi^2$ difference of assigning all tracks to the primary vertex vs. assigning some to a secondary vertex. The SLD algorithm is similar; it searches for points in space from which several tracks originate with high probability. In both algorithms, no attempt is made to form separate $B$ and $D$ vertices.

SLD finds a suitable vertex in $\sim 50\%$ of $b$ hemispheres, compared to $\sim 15\%$ for charm and $\sim 3\%$ for light quarks. A vertex axis is formed by the direction between the primary and this secondary vertex. A final pass is made to assign tracks to the vertex based on their impact parameter to the vertex axis, and the distance along the vertex axis of the closest approach position. Finally, the quantity $M_{p_t} = \sqrt{M^2 + p_t^2} + |p_t|$ is formed, where $M$ is the mass of the tracks in the vertex ( assuming they are all pions) and $p_t$ is their total momentum

Table 14: Fit results and dominant systematic errors for the $D^*l$ analyses.

| Experiment | Dominant Systematic | $\delta(\Delta m_d)$ (ps$^{-1}$) | $\Delta m_d$ (ps$^{-1}$) |
|---|---|---|---|
| CDF | $B^-$ frac. (30%) $D^{**} \to D^*$ frac. (50%) | $\pm 0.020$ $\pm 0.027$ | $0.516 \pm 0.099 \pm 0.035$ |
| DELPHI | $B^-$ frac. (25%) | $\pm 0.012$ | $0.499 \pm 0.053 \pm 0.015$ |
| OPAL | $B^-$ frac. (50%) | $\pm 0.019$ | $0.539 \pm 0.060 \pm 0.024$ |

transverse to the vertex axis. Selected events are required to have $M_{p_t} < 2.0$ GeV, which rejects most of the remaining charm and $uds$ events, leaving 16,803 vertexes with an estimated $b$ purity of 93%.

ALEPH uses the vertex as a seed and accepts tracks with $R_{sig} = S_p/(S_p + S_s) > 0.7$, where $S_{p,s}$ is the significance of the impact parameter of the track to the primary and secondary vertices respectively. Hemispheres with a vertex $\chi^2$ probability $> 1\%$ are accepted – if both hemispheres have a valid vertex, the one with the highest probability is selected as the vertex side of the event.

Table 15: Overview of kaon, dipole and double jet-charge analyses.

| Experiment | Events | Decay Tag | Production Tag |
|---|---|---|---|
| ALEPH [89] | 423,169 | Jet Q | Jet Q$^\dagger$ |
| SLD [90](Kaon) | 5,694 | Kaon | Jet Q + Polarization$^\dagger$ |
| SLD [90](Dipole) | 3,291 | Dipole | Jet Q + Polarization$^\dagger$ |

$^\dagger$ See Table 9.

### 4.6.2   INITIAL- AND FINAL-STATE FLAVOR TAGGING

SLD determines the $b$ flavor of the initial state by using both the jet charge and polarization methods, as described above.

For the final-state tag, two methods were developed. First, the standard $b \to c \to s$ decay chain suggests that inclusive kaons should dominantly have the same sign charge as the $b$ quark. This correspondence is diluted by $B \to D\overline{D}$ decays and by strange quark pairs produced during fragmentation.

Kaons are required to be separated by more than $2.4\sigma$ from the pion hypothesis and by $3.2\sigma$ from a proton. The final-state tag is simply the sum of the charges of all kaon candidates, and so $\sum Q_K > 0$ would tag a $B^0$, for example. The 12% of events with $\sum Q_K = 0$ are discarded. The probability of correctly tagging the charge of the $b$ at its decay is 77% with this method.

The second method involves constructing a charge dipole from the $B$ and $D$ meson decay products. SLD requires the total charge of tracks associated with the vertex to be zero, to enhance the $B^0$ content of the sample as described previously. The direction between the primary and secondary vertices defines the vertex axis. A dipole is formed by summing over positive and negative tracks:

$$\delta q \;=\; \frac{\sum_{i=+} w_i L_i}{\sum_{i=+} w_i} - \frac{\sum_{i=-} w_i L_i}{\sum_{i=-} w_i},$$

where $L_i$ is the location of the distance of closest approach of track $i$ to the vertex axis, and the weight $w_i = \sin^2 \theta_i / \sigma_{Ti}$, where $\theta_i$ is the angle of the track relative to the vertex axis and $\sigma_{Ti}$ is the uncertainty on the impact parameter to the vertex axis. A $B^0 \to l^+ \nu D^-$ decay, for example, will thus tend to have a negative dipole, whereas the equivalent $\overline{B^0}$ decay has a positive dipole. The probability of a correct tag is parameterized as a function of $|\delta q|$, and reaches 68% for large $|\delta q|$.

ALEPH tags both hemispheres using the jet-charge algorithm previously described, using $Q_H(\kappa = 0.5)$ for the opposite side hemisphere, and $Q_H(\kappa = 1.0)$ for the vertex side.

### 4.6.3 FIT METHOD AND RESULTS

Both SLD analyses perform $\chi^2$ fits to the mixed/total (i.e. $N_{DP-}/N$) fraction of the data, binned in decay length. Each event's contribution is weighted by its estimated dilution to maximize the sensitivity (see Section 3.3).

ALEPH converts the decay length to proper time using an event-by-event estimate of the momentum constructed from a weighted sum of the momentum of the charged tracks, energy from neutral clusters associated with the $b$ jet (projected onto the vertex direction) and a correction due to the missing neutrino. They then fit to the weighted charge correlation $< -Q_H^{\mathrm{opp}}(0.5)Q_H^{\mathrm{vtx}}(1.0) > (t)$, i.e. the usual correlation function given by the signs of the hemisphere jet charges, with an event weight given by the product of their magnitudes similarly to the $l$-jet charge analysis described previously.

The results are summarized in Table 16 and Figure 12.

Table 16: Fit results and dominant systematic error for double jet charge, dipole and kaon tag analyses.

| Experiment | Dominant Systematic | $\delta(\Delta m_d)$ (ps$^{-1}$) | $\Delta m_d$ (ps$^{-1}$) |
|---|---|---|---|
| ALEPH | Frag. params. <br> $\Lambda_b$ frac. | $\pm 0.015$ <br> $\pm 0.012$ | $0.441 \pm 0.026 \pm 0.029$ |
| SLD(Kaon) | Fit method <br> Kaon ID efficiency | $\pm 0.042$ <br> $\pm 0.035$ | $0.580 \pm 0.066 \pm 0.075$ |
| SLD(Dipole) | Fit method <br> MC statistics | $\pm 0.016$ <br> $\pm 0.021$ | $0.561 \pm 0.078 \pm 0.039$ |

## 4.7 Same-side Tag Analyses

As mentioned above, in an experiment at a hadron collider such as CDF, it is quite common that one $b$ jet falls within the tracking acceptance while the other $b$ jet does not. Hence a flavor tagging method which only relies on finding one $b$ jet is attractive. Such methods are generally called same-side tags (SST).

### 4.7.1   EVENT SELECTION

The only $b$ samples available to CDF involve semileptonic $b$ decays, selected by either the low-$p_t$ dilepton or high-$p_t$ single lepton triggers described above. Since this analysis is designed to only require one $b$, the single lepton samples are chosen as the basis. The SST algorithm used by CDF [91] requires a relatively pure $B^0$ meson sample, as discussed below. To obtain this, a reconstructed $D^{*+}$ or $D^+$ associated with the lepton is required. Three $D^0$ decay modes, $K^-\pi^+, K^-\pi^+(\pi^0)$, and $K^-\pi^+\pi^+\pi^-$, (with the $\pi^0$ not reconstructed), are used, and the $D^{*+}$ reconstruction follows the standard methods described in Section 4.4. The $D^+$ is reconstructed in the $D^+ \to K^-\pi^+\pi^+$ mode. In addition to the $B^0 \to D^{(*)+}l^-\nu X$ mode, the $B^- \to D^0 l^-\nu X$ mode, with $D^0 \to K^+\pi^-$, is searched for, where the $D^0$ is required to not come from a $D^{*+}$. This signal is used to help estimate the amount of $B^+$ contamination in the (mostly) $B^0$ sample.

### 4.7.2   PRODUCTION FLAVOR TAG

As described by the schematic picture in Figure 8, the SST tag used by CDF attempts to find the "first" charged pion in the fragmentation chain of the $b$ quark. If the $b$ quark forms a $\overline{B^0}$, the first pion should be a $\pi^-$; conversely, a $\pi^+$ would come from a $\overline{b}$ quark. However, if the $b$ quark forms a $B^-$, the pion should be a $\pi^+$, with a $\pi^-$ associated with a $\overline{b}$, exactly the opposite correlation to the $B^0$ case. Hence the charge of the pion, if it can be identified, determines the sign of the $b$ quark, as long as it is known whether the decaying meson is a $B^0/\overline{B^0}$ or $B^\pm$.

The SST pion is a fragmentation track, and so should originate at the primary vertex, not the $B$ decay vertex. SST candidates are thus searched for near the $B$ direction, defined as the $l + D$ direction, with impact parameter significance to the primary vertex of less than three. String fragmentation models suggest that the fragmentation particles have little momentum transverse to the $b$-quark direction, and so the track with the smallest $p_t^{\mathrm{rel}}$ is chosen, where $p_t^{\mathrm{rel}}$ is the track's momentum orthogonal to the direction given by the $l + D$+SST pion system. A suitable candidate can be found in approximately 70% of the events.

Table 17: Event selection and same-side tag properties from CDF.

| | $B^0$ sample, $l^-D^{(*)+}$ $(D^{*+} \to D^0\pi^+_*)$ | | $B^-$ sample, $l^-D^0$ | |
|---|---|---|---|---|
| Modes/Events | $D^0 \to K^-\pi^+$ $K^-\pi^+\pi^+\pi^-$ $D^0 \to K^-\pi^+(\pi^0)$ $D^+ \to K^-\pi^+\pi^+$ | 1754 2515 1997 | $D^0 \to K^-\pi^+$ (not from $D^{*+}$) | 2928 |
| Tag Effic. | 70% | | 70% | |
| Tag $<P_r>$ | 0.59 | | 0.64 | |

### 4.7.3   FIT METHOD AND RESULTS

The decay length and conversion to proper decay time is performed as in the $D^*l - l$ analysis described in Section 4.5.

The sign of the lepton in the $lD^{(*)}$ decay tags the $b$ quark charge at its decay point. For the predominantly $B^0$ sample, given by the $lD^{*+}$ events, the "unmixed" $DP+$ sample has same-sign, $l^- - \pi_{SST}^-$ and $l^+ - \pi_{SST}^+$ events, whereas the $DP-$ sample has opposite-sign events. For the $B^+$ sample the assignment is reversed.

Each sample is divided into six bins in proper decay length, and the mean value of the charge-correlation function $<q_c(ct)> = [N_{DP+}(ct) - N_{DP-}(ct)]/[N_{DP+}(ct) + N_{DP-}(ct)]$ is evaluated for each bin. If the samples were pure $B^0$ and $B^+$, the first would show the usual $\cos \Delta m_d t$ dependence whereas the second would be flat. However, there is some cross contamination between the samples. First, $B^0$ events with $D^0 \to K^- \pi^+$ decays end up in the $B^+$ sample if the $\pi_*^+$ from the $D^{*+}$ decay is missed. CDF determined that this is the case for $15 \pm 7\%$ of these $B^0$ events. The most serious cross contamination, however, comes from decays involving the P-wave $D^{**}$ resonances (or non-resonant $D^{(*)}\pi$ pairs), as shown in Figure 7. For example, the decay $B^- \to D^{**0}l^-\nu$ followed by $D^{**0} \to D^{*+}\pi_{**}^-$ (denoting the $\pi$ from the $D^{**}$ as $\pi_{**}$) is classified as a $B^0$ decay. A further complication is that the $\pi_{**}$ may be selected as the SST pion. As this example shows, this pion's charge is always correlated with that of the lepton such that the event is classified as $DP+$, which biases the charge-correlation function. The $\pi_{**}$ particles, however, originate from the $B$-decay vertex, whereas the SST pions are selected to come from the primary vertex, which suppresses this effect.

The fraction of $D^{**}$ in semileptonic $B$ decays, $f^{**}$, and the branching fractions of the $D^{**}$ mesons are rather poorly known. CDF assumes $f + f^* + f^{**} = 1$, where $f$ and $f^*$ are the decay fractions into $lD$ and $lD^*$. CDF sets $f^{**} = 0.36 \pm 0.12$ [92] and $f^*/f = 2.5 \pm 0.6$ [93]. Further, there are four $D^{**}$ states of differing spin-parity, some of which decay to $D^*\pi$ and some to $D\pi$. The relative abundances of the $D^{**}$ states in the $B$ decay, then, have an effect on the amount of cross contamination among the samples by changing the fraction involving the $D^*$ state. This fraction is left as a free parameter in the fit and is found to be $0.3 \pm 0.3$.

Table 18: Fit result and dominant systematic errors for the same-side tag analysis.

| Experiment | Dominant Systematic | $\delta(\Delta m_d)$ (ps$^{-1}$) | $\Delta m_d$ (ps$^{-1}$) |
|---|---|---|---|
| CDF [91] | $D^{**}$ fraction(30%) $D^{**} \to D^*$ fraction (100%) | $\pm 0.031$ | $0.471 \pm 0.078 \pm 0.034$ |

### *4.8   Summary of $B^0$ Mixing Results*

As this exhaustive summary of methods and results shows, a large number of different techniques are used to extract the oscillation frequency of $B^0$ mixing. Within each single technique, each experiment typically has unique methods of measuring the decay length and momentum of the $B$ meson. Further, in evalu-

ating systematics, the range of variation of variables, even those common to all analyses, is often not consistent.

Therefore, combining all these results must be done with care. Since there are so many measurements, the final error can quickly become dominated by systematic effects, and the handling of common systematics is especially important. A $B$ Oscillation Working Group was formed [94] to combine the measurements in as consistent a manner as possible.

Their results, shown in Figure 12, give a best fit of $\Delta m_d = 0.484 \pm 0.015$ ps$^{-1}$, including the time-integrated ARGUS and CLEO measurements.

## 5   $B_s^0$ MIXING

That $B_s^0$ mesons undergo oscillations is clear. The time-integrated probability that a $B^0$ mixes, $\chi_d$, is measured by the ARGUS and CLEO experiments, which operate at $e^+e^-$ colliders tuned to the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ is not sufficiently heavy to decay to $B_s^0\overline{B_s^0}$ pairs, and so the only contribution to a mixing signal is from $B^0$ mesons. At LEP, SLD and CDF, however, both $B^0$ and $B_s^0$ are produced, so time-integrated mixing measurements determine $\overline{\chi} = f_d\chi_d + f_s\chi_s$, where $f_d \approx 40\%$ is the $B^0$ fraction in the $b$ sample, and $f_s \approx 11\%$ is the $B_s^0$ fraction. Since $\chi = 0.172 \pm 0.010$ and $\overline{\chi} = 0.118 \pm 0.006$ [6], a large value for $\chi_s$ is required, indicating that $B_s^0$ mesons do indeed mix.

The problem is that $B_s^0$ mesons oscillate at a high frequency, so $\chi_s \approx 0.5$ and $\chi_s$ has no power to resolve $\Delta m_s$ near this limit. Unfortunately, no experiment has yet succeeded in directly measuring the frequency of $B_s^0$ oscillations; only lower limits have been determined.

For several reasons, including the notorious difficulty in combining exlusion regions from several experiments, an alternate method of fitting for $\Delta m_s$ was proposed by Moser and Roussarie [95]. This method, commonly called the amplitude method, provides a consistent fitting procedure for all experiments, making it much easier to combine the results. Essentially this method entails searching for a peak in the power spectrum of the data as a function of frequency. The likelihood function for a given data sample is constructed in the usual way, by estimating the amounts of various components of the sample. Each component's expected proper-time distribution is formed based on its true distribution, convoluted with the decay-length and momentum-resolution functions, and combined in the appropriate fractions. The amplitude method introduces one difference in the construction of the likelihood function: the true probability densities of the $B_s^0$ and $\overline{B_s^0}$ mesons are replaced by

$$P(t) \;=\; \frac{\Gamma_s}{2}e^{-\Gamma_s t}(1 \pm \mathcal{A}\cos\Delta m_s t), \tag{34}$$

i.e. an amplitude $\mathcal{A}$ is introduced in front of the oscillation term.

The new likelihood function is then maximized as a function of $\mathcal{A}$ alone, i.e. all parameters, such as sample composition parameters, $B$ hadron lifetimes, and even $\Delta m_s$, are fixed. This gives a measure of $\mathcal{A}(\Delta m_s)$ and an error $\sigma_{\mathcal{A}}(\Delta m_s)$ for that value of $\Delta m_s$. This fit is repeated for many different values of $\Delta m_s$ to sketch out the shape of the $\mathcal{A}(\Delta m_s)$ curve. If $\Delta m_s = \Delta m_s^{\text{true}}$, the fit should return $\mathcal{A} = 1$, within its errors. For $\Delta m_s$ far from its true value, $\mathcal{A}$ should fluctuate around 0, within its errors.

A given $\Delta m_s$ can thus be excluded at greater than 95% CL if $\mathcal{A}+1.645\sigma_{\mathcal{A}} \leq 1$,
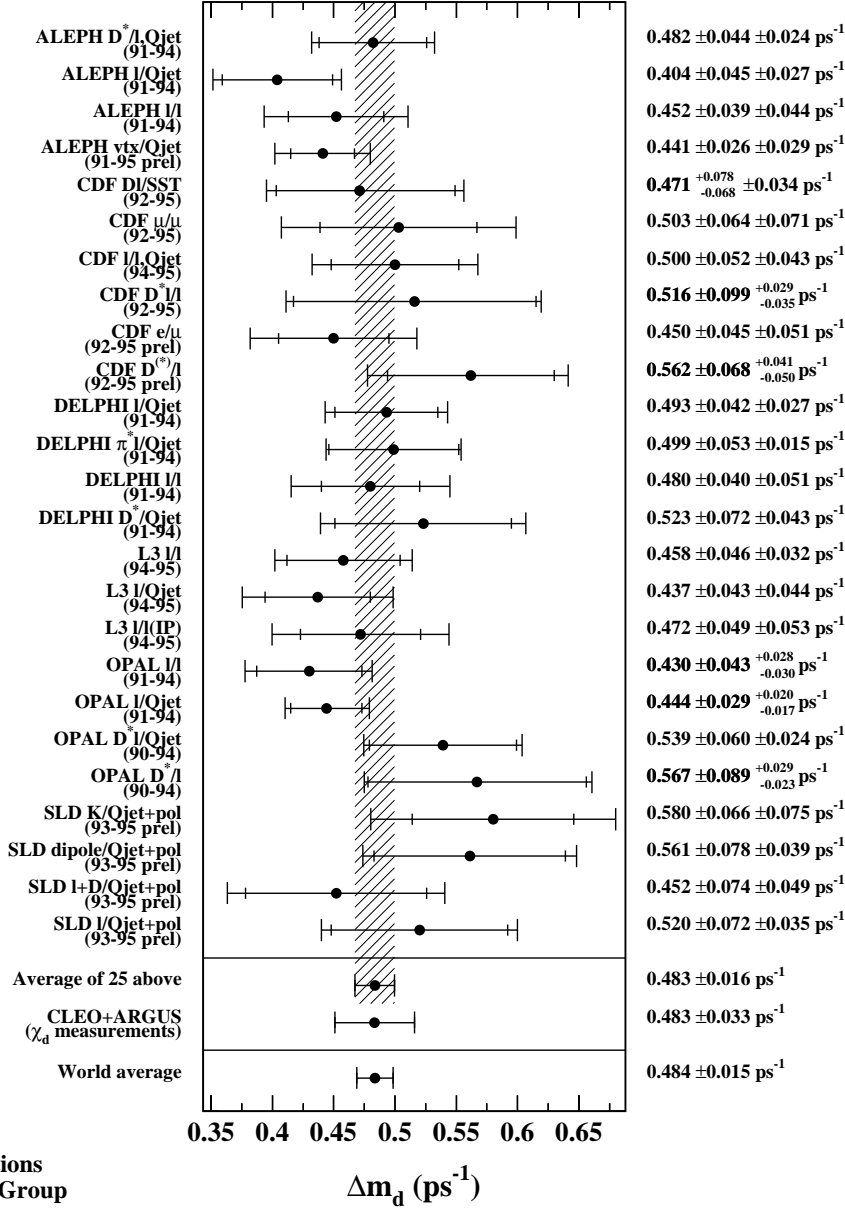
Figure 12: Summary of $B^0$ mixing results from the $B$ Oscillation Working Group [94].

assuming $\sigma$ represents a gaussian error. An experiment's 95% CL lower limit on $\Delta m_s$ is defined as the largest $\Delta m_s$ such that all smaller values have measurement probabilities below 5%.

Of course, as with all limits, for a given error $\sigma_{\mathcal{A}}$, fluctuations in the central value $\mathcal{A}$ can result in more or less favourable exlusion regions. The sensitivity of an experiment is therefore defined as the largest $\Delta m_s$ excluded if $\mathcal{A} = 0$ for all

$\Delta m_s$, i.e. the point at which $1.645\sigma_{\mathcal{A}}(\Delta m_s) = 1$.

The lower limits on $\Delta m_s$ and sensitivities of the analyses presented here are dominated by their statistical errors. The values quoted include a degradation due to systematic effects, but since these change the result very little, the major systematic errors are not listed here.

Searches for $B_s^0$ mixing fall into two rough categories: high statistics but low $B_s^0$-purity analyses, based primarily on inclusive lepton samples, with typical $B_s^0$ purities of $\sim 10\%$; and low statistics, high-purity ($\sim 60\%$) samples in which the $D_s$ daughter of the $B_s^0$ has been reconstructed. In order to obtain the maximal sensitivity to $\Delta m_s$, these experiments combine several production flavor tags.

### 5.1 Inclusive Analyses

These analyses closely follow the form of their $B^0$ counterpart – the inclusive $l$-jet charge analyses described in Section 4.3.

### 5.1.1 EVENT SELECTION AND VERTEXING

ALEPH [96], DELPHI [97], OPAL [98] and SLD [99, 100] (with two analyses, referred to here as $l$-track and $lD$) search for $B_s$ oscillations in the inclusive lepton samples described in Section 4.3. A third SLD analysis (referred to as "dipole") is based on an inclusive vertex sample similar to the one used for the dipole-based $B^0$ analysis described in Section 4.6. Table 19 shows an overview of the event selection for these analyses.

Table 19: Event selection for the inclusive-type $B_s^0$ mixing measurements. All analyses except the SLD dipole are based on single lepton or dilepton samples.

| Experiment | $p(l)$ (GeV) | $p_t^{\text{rel}}(l)$ (GeV) | Events | $B_s^0$ fraction |
|---|---|---|---|---|
| ALEPH [96] | 3.0 | 1.25 | 33,023 | 10.4% |
| DELPHI [97] | 3.0 | 1.0 | 78,476 | 10.5% |
| OPAL [98] | 2.0 | 0.7 | 53,050 | 10.5% |
| SLD [99] ($l$+track) | $1.0(e), 2.0(\mu)$ | 0.8 | 9,691 | 8.5% |
| SLD [100] ($l$+D) | | 0.9 | 2,009 | 15.9% |
| SLD [100] (dipole) | N/A | N/A | 7,547 | 15.6% |

ALEPH, DELPHI and SLD ($l$-track and $lD$ analyses) search for a vertex by applying the same inclusive algorithms used in the $B^0$ $l$-jet charge analyses, described in Section 4.2 (ALEPH and DELPHI) and Section 4.3 (SLD). For the SLD $lD$ sample, a $D$-like vertex is required in the events. Demanding the total charge of tracks coming from the $B$ and $D$ vertices to be zero enhances the $B_s^0$ (and $B^0$) content of this sample to 15.9%. DELPHI and SLD reconstruct the $B$ decay length and momentum following the procedure used for the $B^0$ analyses; in addition, SLD adds a few extra requirements designed to improve the decay length resolution and supress backgrounds from cascade $b \to c \to l$ decays (such as the increased $p_t^{\text{rel}}$ requirement).

ALEPH adds requirements on the angle that the charm pseudotrack (produced

by the vertexing algorithm) makes with the lepton and with the jet, and on the mass of the $B$ decay products. These cuts, which are not present in the $B^0$ analysis, reduce the number of events by $\sim 65\%$, but decrease the non-$b$ background by a factor of $\sim 4$ and significantly increase both the decay-length and momentum resolution of the remaining events. The energy of the charm particle is estimated by clustering charged and neutral particles with energy $> 0.5$ GeV (to reduce fragmentation clutter) with the charm vertex's tracks until the mass exceeds 2.7 GeV. The neutrino energy is estimated from the missing energy in the lepton hemisphere using the beam constraint as described above. The $B$ momentum is then estimated by $p_B = \sqrt{(E_c + E_\nu + E_l)^2 - m_B^2}$. The core resolutions given by these methods are $280\mu$m on the decay length and $7\%$ on the fractional momentum resolution.

OPAL uses a different vertexing algorithm than that used in the $B^0$ inclusive lepton analysis. They attempt to find both the $b$ and charm decay vertex simultaneously. They form a likelihood to which each track contributes $(w/2) \times (P_b + P_c) + (1 - w) \times P_p)$, where $P_p, P_b$ and $P_c$ are the probabilities that the track originates from the primary, $b$ or charm vertex, as determined from the track's (three-dimensional) impact parameters to these vertex positions, and $w$ is the probability that the track originates from the $b$ vertex, determined from the track momentum and angle relative to the $b$-jet direction. For the lepton track, $w$ is set to 1.

The inclusive vertexing algorithm used in SLD's dipole analysis differs from that used in the $B^0$ analysis. In brief, it relies upon the fact that the $D$ and $B$ flight directions are almost identical. This implies that a single straight line (in three-dimensional space) can be found such that all the $B$ and $D$ decay tracks intersect it at one of two points. The algorithm first tries to find this line by minimizing a $\chi^2$ based on the intersections of all charged tracks with the line. Once this line is found, it is assigned a finite width and considered a "ghost" track. Tracks are again vertexed with this track or the primary vertex to build up the primary, secondary ($B$) and tertiary ($D$) vertices (see Reference [100] for the details of this method). The $B$ momentum is estimated as in the $lD$ analysis.

### 5.1.2   FLAVOR TAGGING

All the inclusive lepton analyses use the sign of the lepton to tag the $b$ flavor at the decay point . The SLD dipole analysis defines a charge dipole $\delta Q \equiv l_{BD} \times \text{sign}(Q_D - Q_B)$, where $Q_B$ and $Q_D$ are the charges of the $B$ and $D$ vertices, and $l_{BD}$ is the distance between them. As in the SLD $B^0$ dipole analysis, $\delta Q > 0$ tags a $\overline{B}$.

ALEPH uses a combination of three production flavor tags:

- Opposite side lepton ($p > 3.0$ GeV) and opposite-side jet charge ($Q_H(0.5)$). See Sections 4.2 and 4.3. No $p_t^{\text{rel}}$ cut is imposed on the lepton.

- Fragmentation (SST) kaon: The kaon candidate, selected by dE/dx, is required to be more consistent with originating from the primary vertex than the secondary vertex and have direction within $45°$ of the $B_s^0$ direction.

The data are categorized according to the presence of these tags. The lepton tag always takes precedence if it exists. If not, the kaon tag then used, if present; otherwise the sign of the jet-charge serves as the tag. The average right-sign probability of this tag is $71\%$.

DELPHI uses the following combination of production tag variables in the hemisphere opposite the lepton:

- The opposite-side jet charge, $Q_H(0.6)$ (the hemisphere jet charge is defined in Section 4.3); the opposite-side kaon charge, $Q_K$ (the jet charge evaluated for kaon candidates alone); and the opposite-side lepton, weighted by its $p_t^{\text{rel}}$ (see Section 4.2).
- The sum of charges of tracks not compatible with the primary event vertex and the sum of charges of tracks compatible with the primary event vertex

In the lepton hemisphere, DELPHI uses:

- The vertex-side jet charge: $Q_H(0.6)$
- The fragmentation (SST) kaon or fragmentation (SST) $\Lambda^0$ (the rapidity of the highest-momentum kaon or $\Lambda^0$ candidate relative to the thrust axis that is compatible with being a fragmentation track, signed by its charge.)

All this information, along with a tag based on the polar angle of the thrust axis (see Section 3.2), is combined into one variable. The average right-sign probability of this tag is 69%.

OPAL also uses a combination of tags:

- The opposite-side lepton – a lepton passing the standard OPAL requirements.

- The modified jet charge – a combination of jet charges from both the opposite and vertex hemisphere, as well as fragmentation kaon information and opposite-side vertex charge. On the vertex side, $Q_{\text{vtx}}$ combines the unweighted $b$-jet charge [$Q_H(0)$, with $Q_H$ given in Table 9] and the hemisphere charge [$Q_H(0)$ summed over all tracks in the hemisphere], as well as two other jet charges (with track weights given by neural nets which were fed the track's kaon and pion probabilities and momentum, direction and impact parameter information.
  The opposite hemisphere's tag, $Q_{\text{opp}}$, combines the standard jet-charges $Q_H(1.0)$ and $Q_H(0)$, the output of a neural net similar to that used in the vertex hemisphere, and, if available, the total charge of tracks in a viable secondary vertex.
  The tag variables $Q_{\text{vtx}}$ and $Q_{\text{opp}}$ are used to form the final variable

$$Q_{\text{tag}} \quad = \quad \frac{2Q_{\text{vtx}}Q_{\text{opp}}}{Q_{\text{vtx}}Q_{\text{opp}} + (1 - Q_{\text{vtx}}Q_{\text{opp}})} - 1.$$

OPAL estimates this tag variable is 40% more effective than the combined jet charge used in the $B^0$ $l$-jet charge analysis.

All three SLD analyses use the same combination of jet-charge and polarization (polar angle) initial-state tags described in the $B^0$ $l$-jet charge section.

## 5.1.3   FIT METHOD AND RESULTS

The sample composition of the lepton hemisphere is determined event-by-event. It is similar to that of the inclusive lepton samples described in Section 4.3. The resulting increase in sensitivity is equivalent to having approximately 30% more events. All experiments perform amplitude fits to the data. The individual lower limits on $\Delta m_s$ and analysis sensitivities are listed in Table 22.

## 5.2 Analyses with Reconstructed $D_s$ Decays

The remaining analyses take a more exlusive approach, ranging from partially reconstructing the $D_s$ decay in $B_s^0 \to D_s l$ decays, to fully reconstructing the $D_s$ decay but partially reconstructing the $B_s^0$ in $B_s^0 \to D_s l X, D_s h X$ decays, to fully reconstructing the $B_s^0$ decay. Clearly the latter will give ideal vertex and momentum determinations, since all the $B_s^0$ decay products are found; however the very small exclusive $B_s^0$ branching ratios mean that the event sample is rather small.

ALEPH and DELPHI reconstruct many $D_s$ decay modes, summarized in Table 20. The strange mesons are reconstructed in their usual accessible modes: $\phi \to K^+ K^-$, $\overline{K}^{*0} \to K^- \pi^+$, $\overline{K}^{*+} \to K_s^0 \pi^+$, and the $f$ resonance as $f(980) \to \pi^+ \pi^-$. For details on the selection criteria, see the references indicated in Table 21.

Table 20: $D_s$ decay modes searched for by various experiments.

| $D_s$ mode | ALEPH | | DELPHI | | | CDF |
| | $D_s l$ | $D_s h$ | $D_s l / \phi l$ | $D_s h$ | Full $B_s^0$ | $\phi l$ |
| --- | --- | --- | --- | --- | --- | --- |
| $\phi \pi^+$ | Y | Y | Y | Y | Y | |
| $\phi \pi^+ \pi^0$ | Y | | Y | | | |
| $\phi \pi^+ \pi^- \pi^+$ | Y | | Y | | Y | |
| $\overline{K}^{*0} K^+$ | Y | Y | Y | Y | Y | |
| $\overline{K}^{*0} K^{*+}$ | Y | | Y | | Y | |
| $K_s^0 K^+$ | Y | Y | Y | | Y | |
| $f(980) \pi^+$ | | | | | Y | |
| $\phi l \nu_l (l = e, \mu)$ | Y | Y | Y | | | |
| $\phi h^+$ | | | Y | | | Y |

### 5.2.1 EVENT SELECTION

ALEPH [101, 102] and DELPHI [103, 104] reconstruct $B_s^0 \to D_s^- l^+ X$ and $B_s^0 \to D_s^- h^+ X$, with the $D_s$ decaying in the modes given in Table 20. In addition, DELPHI searches for fully reconstructed $B_s^0$ (up to a missing $\gamma$ or $\pi^0$) in the following modes: $D_s^- \pi^+, D_s^- a_1^+, \overline{D^0} K^- \pi^+$, and $\overline{D^0} K^- a_1^+$, with $a_1^+ \to \rho^0 \pi^+ \to \pi^+ \pi^- \pi^+$ and $\overline{D^0} \to K^+ \pi^-, K^+ \pi^- \pi^+ \pi^-$.

Note that for $B_s^0 \to D_s^* \to D_s \gamma, D_s \pi^0$ transitions, the photon or $\pi^0$ may not be reconstructed. Thus the mass spectrum of the $B_s^0$ candidates is expected to have a sharp peak at the $B_s^0$ mass, and a wider satellite peak at lower mass from $D_S^*$ decays. DELPHI finds $8 \pm 4$ signal events out of the 11 events in the $B_s^0$ mass region, and $15 \pm 8$ signal events out of 33 in the satellite region.

In addition to the exclusive $D_s$ modes listed, for events with a lepton, DELPHI also searches for a partially reconstructed $D_s$ of the form $D_s \to \phi h^+ X$ (which was not found as one of the above modes).

CDF [105] searches for $\phi \to K^+ K^-$ candidates in the low-$p_t$ dilepton samples ($\mu\mu$ and $e\mu$) described in Section 4.2. The lepton on the $\phi$ side is required to have $p_t^{\text{rel}} > 1.0$ GeV. A charged hadron near the $\phi l$ pair is required, and the effective

Table 21: Summary of the $B_s$ decay modes, sample sizes and purity, and right-sign tag probability for analyses with a partially or fully reconstructed $D_s$.

| Experiment | $B_s$ Mode | Events | $B_s^0$ frac. | $<P_r>$ |
|---|---|---|---|---|
| ALEPH [101, 102] | $D_s l$ | 277 | 66% | 0.73 |
|  | $D_s h$ | 1,620 | 22% | 0.74 |
| CDF [105] | $\phi h l$ | 1,068 | 61% | 0.76 |
| DELPHI [103, 104] | $D_s l$ | 436 | 53% | 0.78 |
|  | $\phi h l$ | 441 | 9% | 0.78 |
|  | $D_s h$ | 2,953 | 20% | 0.78 |
|  | $B_s^0$ (full) | 44 | 51% | 0.78 |

masses are restricted to be $1.0 < m_{\phi h} < 2.0$ and $m_{\phi h l} < 5.0$, consistent with decay kinematics for $B_s^0 \to D_s l \nu$ and $D_s \to \phi h X$.

Table 21 gives an overview of these data samples.

### 5.2.2   DECAY LENGTH AND PROPER TIME

A $B$ decay vertex is easily obtained in these events. First the $D_s$ decay products are vertexed and a pseudotrack formed, which passes through this vertex and has a momentum given by the $D_s$ momentum. This track is intersected with the other track available, $l$ or hadron, to give the $B_s^0$ decay point.

For the $D_s l$ samples, ALEPH and DELPHI estimate the $B$ momentum based on the momentum of the $D_s l$ system and apply a correction for the neutrino obtained from the missing energy in the hemisphere. For the $D_s h$ sample, ALEPH adds charged and neutral particles to the $D_s h$ system and selects the most likely combination based on the effective mass. The momentum of the fully reconstructed decay is simply measured from all the decay products. CDF corrects the measured momentum of the $\phi h l$ system for the missing energy using a mean value derived from Monte Carlo simulation.

### 5.2.3   FLAVOR TAGGING

All experiments use the sign of the lepton on the vertex side as the final-state flavor tag. ALEPH and DELPHI use the combination of production flavor tags described in the inclusive lepton section above, whereas CDF uses the sign of the opposite-side lepton.

### 5.2.4   FIT METHOD AND RESULTS

DELPHI estimates the right-tag probability event-by-event by using the shape of the $P_r(x)$ distributions, which are determined from Monte Carlo simulation. This improves the effective mean right-tag probability, e.g., from 0.745 to 0.78 for the 1994-1995 $D_s l$ data. The effectiveness of the combined tags can be seen by comparing the right-sign tag probabilities in Table 21 to the individual tags used in the $B^0$ analyses.

All the experiments perform an amplitude style fit, with the $\Delta m_s$ exclusion regions and sensitivities given in Table 22.

*5.3  Summary of $B_s^0$ Mixing*

Table 22: Excluded regions of $\Delta m_s$ and the sensitivity of each analysis.

| Experiment | Excluded $\Delta m_s$ region (95% CL) | Sensitivity |
|---|:---:|:---:|
| ALEPH $l$ [96] | $< 9.5$ ps$^{-1}$ | 9.5 ps$^{-1}$ |
| ALEPH $D_s l$ [101] | $< 6.6$ ps$^{-1}$ | 6.7 ps$^{-1}$ |
| ALEPH $D_s h$ [102] | $< 3.9$ ps$^{-1}$, $6.5 - 8.8$ ps$^{-1}$ | 4.1 ps$^{-1}$ |
| CDF $\phi l - l$ [105] | $< 5.8$ ps$^{-1}$ | 5.1 ps$^{-1}$ |
| DELPHI $l$ [97] | $< 4.7$ ps$^{-1}$ | 6.5 ps$^{-1}$ |
| DELPHI $D_s l$ [103] | $< 7.5$ ps$^{-1}$ | 8.2 ps$^{-1}$ |
| DELPHI $B_s^0 + D_s h$ [104] | $< 4.0$ ps$^{-1}$ | 3.2 ps$^{-1}$ |
| OPAL $l$ [98] | $< 5.2$ ps$^{-1}$ | 7.2 ps$^{-1}$ |
| SLD $l$ [99] | $< 1.3$ ps$^{-1}$, $2.0 - 8.6$ ps$^{-1}$, $9.8 - 12.2$ ps$^{-1}$ | 3.8 ps$^{-1}$ |
| SLD $lD$ [100] | $< 5.2$ ps$^{-1}$ | 3.5 ps$^{-1}$ |
| SLD dipole [100] | $< 5.2$ ps$^{-1}$ | 5.4 ps$^{-1}$ |

Table 22 summarizes the excluded regions of $\Delta m_s$ from each of the analyses presented, along with their sensitivities. The *B* Oscillation Working Group has combined these, and finds that $\Delta m_s > 14.6$ ps$^{-1}$ at the 95% CL [94]. Figure 13 contains the combined amplitude plot showing the value of the amplitude $\mathcal{A}$ in Equation 34 versus assumed $\Delta m_s$, from which the lower limit and experimental sensitivity can be read.

# 6  $B_s^0$ LIFETIME DIFFERENCE

Another, more indirect way of searching for $B_s^0$ mixing utilizes the relationship between the mass difference of the two eigenstates, $\Delta m_s$, and the decay rate difference $\Delta\Gamma_s$ [106],

$$\frac{\Delta\Gamma_s}{\Delta m_s} = -\frac{3}{2}\pi\frac{m_b^2}{m_t^2}\frac{\eta_{\text{QCD}}^{\Delta\Gamma_s}}{\eta_{\text{QCD}}^{\Delta m_s}}. \tag{35}$$

The QCD correction factors $\eta$ have recently been estimated, giving $\Delta\Gamma_s/\Delta m_s = (5.6 \pm 2.6) \times 10^{-3}$ [10].

A large $\Delta m_s$ corresponds to a large lifetime difference between the heavy and light states. Even a moderately large $\Delta m_s \approx 20$ would result in a 17% difference in their lifetimes.

This difference has been searched for in two ways. First, $B_s^0$ decays to final states that are not CP eigenstates should contain a mixture of both $B_H$ and $B_L$ components. Several experiments have fit their $B_s^0$ lifetime distributions to a sum of two exponentials of decay widths $\Gamma_s \pm \Delta\Gamma_s/2$. From their sample of $\sim 600$ $B_s^0 \to l^+ D_s^- \nu X$ events, CDF [107] finds $\Delta\Gamma_s/\Gamma_s = 0.34^{+0.31}_{-0.34}$ (statistical only), corresponding to an upper limit of $\Delta\Gamma_s/\Gamma_s < 0.83$ (95% CL). DELPHI [108], using their $D_s^\pm l^\mp$ and $D_s^\pm h^\mp$ samples from the $B_s^0$ mixing analysis (Section 5.2), finds $\Delta\Gamma_s/\Gamma_s < 0.42$ (95% CL).
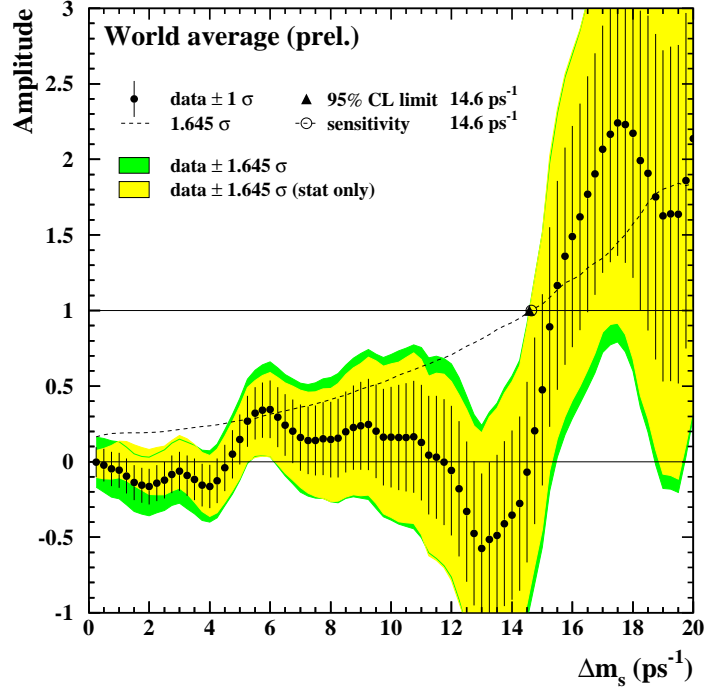
Figure 13: $B$ Oscillation Working Group's [94] combined plot of fitted amplitude versus $\Delta m_s$, for the $B_s$ mixing analyses presented in this section.

L3 [109] fits 445,000 events, selected using an inclusive vertexing algorithm, to a sum of $B^0$, $B^+$, $\Lambda_b$ exponential decay distributions and a two-lifetime $B_s^0$ component, resulting in $\Delta\Gamma_s/\Gamma_s = 0.00^{+0.30}_{-0.00}$ (statistical only). Including systematic effects, L3 estimates an upper limit of $\Delta\Gamma_s/\Gamma_s < 0.67$ (95% CL).

A second method to search for a lifetime difference is to measure the lifetime of $B_s^0$ decays to modes that are CP eigenstates or are mixtures dominanted by one CP state. These decays arise primarily from either $B_H$ or $B_L$ decays, depending on the CP eigenvalue of the final state. The difference between the lifetime measured in this sample and the mean lifetime measured over a mixture of final states measures $\Delta\Gamma$. For example, ALEPH [110] found $32 \pm 17$ decays of $B_s^0 \to D_s^{(*)+} D_s^{(*)-} \to \phi\phi X$, from which they extracted $\tau(B_s) = 1.42 \pm 0.23 \pm 0.16$ ps. Assuming this is a pure CP-even decay, comparison to the mean $B_s^0$ lifetime [6] of $\tau_{B_s} = 1.55$ ps gives $\Delta\Gamma_s/\Gamma_s = 0.17 \pm 0.35$.

CDF measures the lifetime using $58 \pm 12$ decays of $B_s^0 \to J/\psi\phi$ and finds $\tau(B_s) = 1.34 \pm 0.23 \pm 0.05$ ps [111]. This final state is not necessarily a pure CP eigenstate, and using a transversity analysis [112] CDF finds that the CP-odd component represents $23 \pm 19\%$ of the sample. The higher statistics samples soon to be available at the Tevatron will clearly benefit this type of analysis.

The DELPHI result, combined with the mean $B_s^0$ lifetime, gives an upper limit of $\Delta m_s < 48$ ps$^{-1}$ (95% CL).

# 7 CONCLUSION

Mixing between the particle and antiparticle states of the $B^0$ and $B^0_s$ mesons has several consequences. It leads to physical states that are no longer eigenstates of flavor, particle-antiparticle oscillations can generate CP asymmetries between $B$ and $\overline{B}$ decays and their frequency is relatively directly related to CKM matrix elements. The $B$ system is unique in having two such particles readily available, which can be studied and compared.

Several experiments have succeded in observing the time-dependence of $B^0$ flavor oscillations, using a wide variety of techniques. The measurement of the $B^0_s$ oscillation frequency will have to wait for the next generation of experiments – current experiments are only able to place a lower limit. The principle results presented here are: $\Delta m_d = 0.484 \pm 0.015$ ps$^{-1}$ and $\Delta m_s > 14.6$ ps$^{-1}$.

With the results from Section 2.3, $f_{B_d} = 200 \pm 16$ MeV, $\hat{B}^{\mathrm{NLO}}_{B_d} = 1.37 \pm 0.08$, $f_{B_d}/f_{B_s} = 1.15 \pm 0.06$, $B_{B_d}/B_{B_s} = 1.01 \pm 0.01$ and $\overline{m_t}(174) = 167 \pm 5$ GeV, $\eta_B = 0.55 \pm 0.01$, the CKM matrix element $|V_{td}|$ and ratio $|V_{ts}|/|V_{td}|$ can be extracted according to Equation 24,

$$
\begin{aligned}
\Delta m_d &= \left(24.4 \left[\frac{f_{B_d}}{200 \text{ MeV}}\right]^2 \left[\frac{\hat{B}^{\mathrm{NLO}}_{B_d}}{1.37}\right] \pm 5.0\right) |V_{td}|^2 \quad \text{(eV)} \\
\frac{\Delta m_s}{\Delta m_d} &= (1.36 \pm 0.14) \left|\frac{V_{ts}}{V_{td}}\right|^2 ,
\end{aligned}
$$

corresponding to $|V_{td}| = (3.6 \pm 0.4) \times 10^{-3}$ and $|V_{ts}/V_{td}| > 4.7$ (95% CL).

The future direction for $B$ mixing measurements is clear. Measuring $\Delta m_s$, either directly through a time-dependent oscillation analysis, or through a heavy-light lifetime difference is crucial. The prospects for this occurring in the next few years are good, as the upgraded CDF and DØ detectors are expected to provide significantly improved statistics and sensitivity to measure $\Delta m_s$.

It is equally important to make further progress in calculating $f_B$ and $B_B$. The lattice calculations of these constants are maturing rapidly, and with unquenched simulations becoming available, more precise estimates with a full evaluation of systematic uncertainties should be within reach.

These measurements are an important part of exploring the unitarity of the CKM matrix, along with the various $B$-sector CP violation searches which will take place over the next few years. The study of $B$ mixing will surely remain vibrant for years to come.

*Literature Cited*

1. Gell-Mann M, Pais A. *Phys. Rev.* 97:1387 (1955)
2. Lande K, et al *Phys. Rev.* 103:1901 (1956)
3. Albajar C, et al (UA1 Collaboration). *Phys. Lett.* B186:247 (1987)
4. Albrecht H, et al (ARGUS Collaboration). *Phys. Lett.* B192:245 (1987)
5. Artuso M, et al (CLEO Collaboration). *Phys. Rev. Lett.* 62:2233 (1989)
6. Caso C, et al (Particle Data Group) *Eur. Phys. J. C* 63:1 (1998)
7. Albrecht H, et al (ARGUS Collaboration). *Z. Phys. C* 55:357 (1992); Albrecht H, et al (ARGUS Collaboration). *Phys. Lett.* B324:249 (1994); Bartelt J, et al (CLEO Collaboration). *Phys. Rev. Lett.* 71:1680 (1993)
8. Cabbibo N. *Phys. Rev. Lett.* 10:531 (1963); Kobayashi M, Maskawa K *Prog. Theor. Phys.* 49:652 (1973)
9. Wolfenstein L. *Phys. Rev. Lett.* 51:1945 (1984)
10. Beneke M, Buchalla G, Dunietz I. *Phys. Rev. D* 54:4419 (1996)

11.  Grossman L, Lipkin H. *Phys. Rev. D* 55:2760 (1997)
12.  Gaillard M, Lee B. *Phys. Rev. D* 10:897 (1974)
13.  Fujikawa K, Lee B, Sanda A. *Phys. Rev. D* 6:2923 (1972)
14.  Glashow S, Illiopoulos J, Maiani L. *Phys. Rev. D* 2:1285 (1970)
15.  Inami T, Lim CS. *Prog. Theor. Phys.* 65:297 (1981)
16.  Buras A, Jamin M, Weisz P. *Nucl. Phys.* B347:491 (1990)
17.  Buskulic D, et al (ALEPH Collaboration). *Phys. Lett.* B343:444(1995)
18.  Abreu P, et al (DELPHI Collaboration). CERN-EP-99-162 (1999)
19.  Acciarri M, et al (L3 Collaboration). *Phys. Lett.* B396:327 (1997)
20.  Artuso M, et al (CLEO Collaboration). *Phys. Rev. Lett.* 75:785 (1995)
21.  Bigi I. hep-ph/9508408 (1995); Bigi I, Uraltsev N. *Phys. Lett.* B280:271 (1992)
22.  Abe F, et al (CDF Collaboration). *Phys. Rev. D* 57:5382 (1998)
23.  Abe F, et al (CDF Collaboration). *Phys. Rev. D* 58:092002 (1998).
24.  BaBar Collaboration. *The BaBar Physics Book*. SLAC-R-504 (1998)
25.  Neubert M, Schrajda CT. *Nucl. Phys.* B438:238 (1995)
26.  Sheikholeslami B, Wohlert R. *Nucl. Phys.* B259:572 (1985)
27.  Gavela M, et al *Phys. Lett.* B206:113 (1988)
28.  Bernard C, et al *Phys. Rev. D* 38:3540 (1988)
29.  Boucaud P, et al *Phys. Lett.* B220:219 (1989)
30.  Allton C, et al *Nucl. Phys.* B349:504 (1991)
31.  Alexandrou C, et al *Phys. Lett.* B256:60 (1991)
32.  Duncan A, et al *Nucl. Phys. Proc. Supplement* 30:433 (1992)
33.  Abada A, et al *Nucl. Phys.* B376:172 (1992)
34.  Allton C, et al (APE Collaboration). *Phys. Lett.* B326:295 (1994)
35.  Alexandrou C, et al *Nucl. Phys.* B414:815 (1994)
36.  Bernard C, Labrenz J, Soni A. *Phys. Rev. D* 49:2536 (1994)
37.  Duncan A, et al *Phys. Rev. D* 51:5101 (1995)
38.  Baxter R, et al (UKQCD Collaboration). *Phys. Rev. D* 49:1594 (1994)
39.  Duncan A, et al (HEMCGC Collaboration). *Phys. Rev. D* 49:3546 (1994)
40.  Alexandrou C, et al (PCW Collaboration). *Z. Phys. C* 62:659 (1994)
41.  Hashimoto S. *Phys. Rev. D* 50:4639 (1994)
42.  Allton C, et al (APE Collaboration). *Phys. Lett.* B405:133 (1997)
43.  Becerevic D, et al (APE Collaboration). hep-lat/9811003 (1998)
44.  El-Khadra A, et al *Phys. Rev. D* 58:014506 (1998)
45.  Aoiki S, et al (JLQCD Collaboration). *Phys. Rev. Lett.* 80:5711 (1998)
46.  Bernard C, et al (MILC Collaboration). *Phys. Rev. Lett.* 81:4812 (1998)
47.  Ali Khan A, et al (SGO Collaboration). *Phys. Rev. D* 56:7012 (1997)
48.  Ishikawa K, et al *Phys. Rev. D* 56:7028 (1997)
49.  Ali Khan A, et al (GLOK Collaboration). *Phys. Lett.* B427:132 (1998)
50.  Bernard C, et al (MILC Collaboration). hep-lat/9909121 (1999)
51.  Collins S, et al hep-lat/9901001 (1999)
52.  Ali Khan A, et al (CP-PACS Collaboration). hep-lat/9909052 (1999)
53.  Ali Khan A, et al (CP-PACS Collaboration). hep-lat/9911039 (1999)
54.  Lepage G and Mackenzie P.*Phys. Rev. D* 48:2250 (1993); El-Khadra A, Kronfeld A, Mackenzie P.*Phys. Rev. D* 55:3933 (1997)
55.  Bernard C, et al (MILC Collaboration). hep-lat/9709142 (1997)
56.  Albrecht H, et al (Argus Collaboration). *Z. Phys. C* 54:1 (1992); Aoki S, et al (WA75 Collaboration). *Prog. Theor. Phys.* 89:131 (1993); Kodama K, et al (E653 Collaboration). *Phys. Lett.* B382:299 (1996); Bai J, et al (BES Collaboration). *Phys. Rev. Lett.* 74:4599 (1995); Acciari M, et al (L3 Collaboration). *Phys. Lett.* B396:327 (1997); Chadha M, et al (CLEO Collaboration). *Phys. Rev. D* 58:032002 (1998); Abreu P, et al (DELPHI Collaboration). DELPHI 97-105 CONF 87 (1997); Buskulic D, et al (ALEPH Collaboration). ALEPH 98-063 (1998)
57.  Khodjamirian A, Ruckl R. hep-ph/9801443 (1998)
58.  Draper T hep-lat/9810065 (1998)
59.  Gimenez V, Reyes J. hep-lat/9806023 (1998)
60.  Flynn J, Sachrajda CT. hep-lat/97100057 (1997)
61.  Gimenez V, Martinelli G.*Phys. Lett.* B398:135 (1997)
62.  Bernard C, Blum T, Soni A. hep-lat/9801039 (1998)

63. Aoiki S, et al (JLQCD Collaboration). hep-lat/9809152 (1998)

64. Aoiki S, et al (JLQCD Collaboration). hep-lat/9510033 (1995)

65. Ewing A, et al (UKQCD Collaboration). *Phys. Rev. D* 54:3526 (1996)

66. Bernard C, Soni A. *Nucl. Phys.* B47:43 (1996)

67. Acton P, et al (OPAL Collaboration). *Phys. Lett.* B307:247 (1993)

68. Gronau M, Nippe A, Rosner J. *Phys. Rev. D* 47:1988 (1993)

69. Buskulic D, et al (ALEPH Collaboration). *Phys. Lett.* B313:498 (1993); Buskulic D, et al (ALEPH Collaboration). *Phys. Lett.* B322:441 (1994)

70. Buskulic D, et al (ALEPH Collaboration). *Nucl. Instr. Meth. A* 360:481 (1995)

71. Abe F, et al (CDF Collaboration). *Nucl. Instr. Meth. A* 271:387 (1988); Abe F, et al (CDF Collaboration). *Phys. Rev. D* 50:2966 (1994)

72. Abreu P, et al (DELPHI Collaboration). *Nucl. Instr. Meth. A* 378:57 (1996)

73. Adeva B, et al (L3 Collaboration). *Nucl. Instr. Meth. A* 289:35 (1990); Adriani O, et al (L3 Collaboration). *Phys. Rep.* 236:1 (1993)

74. Ahmet K, et al (OPAL Collaboration). *Nucl. Instr. Meth. A* 305:275 (1991); Allport P, et al (OPAL Collaboration). *Nucl. Instr. Meth. A* 324:34 (1993); Allport P, et al (OPAL Collaboration). *Nucl. Instr. Meth. A* 346:476 (1994)

75. Abe K, et al (SLD Collaboration). *Phys. Rev. D* 53:1023 (1996); Abe K, et al (SLD Collaboration). *Nucl. Instr. Meth. A* 400:287 (1997)

76. Buskulic D, et al (ALEPH Collaboration). *Z. Phys. C* 75:397 (1997)

77. Abe F, et al (CDF Collaboration). *Phys. Rev. D* 60:051101 (1999)

78. Abe F, et al (CDF Collaboration). *Phys. Rev. D* 60:072003 (1999)

79. Abe F, et al (CDF Collaboration). CDFNOTE 3791, Preliminary (1996)

80. Abreu P, et al (DELPHI Collaboration). *Z. Phys. C* 76:579 (1997)

81. Acciarri M, et al (L3 Collaboration). *Eur. Phys. J. C* 5:195 (1998)

82. Ackerstaff K, et al (OPAL Collaboration). *Z. Phys. C* 76:417 (1997)

83. Ackerstaff K, et al (OPAL Collaboration). *Z. Phys. C* 76:401 (1997)

84. Abe K, et al (SLD Collaboration). SLAC-PUB-7228 (1996)

85. Abe K, et al (SLD Collaboration). SLAC-PUB-7229 (1996)

86. Abe F, et al (CDF Collaboration). CDFNOTE 4526, Preliminary (1998)

87. Alexander G, et al (OPAL Collaboration). *Z. Phys. C* 72:377 (1996)

88. Affolder T, et al (CDF Collaboration). *Phys. Rev. D* 60:112004 (1999)

89. Buskulic D, et al (ALEPH Collaboration). EPS-HEP Jerusalem, Contribution 596 (1997)

90. Abe K, et al (SLD Collaboration). SLAC-PUB-7230 (1996)

91. Abe F, et al (CDF Collaboration). *Phys. Rev. D* 59:032001 (1999)

92. Fulton R, et al (CLEO Collaboration). *Phys. Rev. D* 43:1 (1991)

93. Barnett R, et al (Particle Data Group). *Phys. Rev. D* 54:1 (1996)

94. *B Oscillation Working Group.* *http://lepbosc.web.cern.ch/LEPBOSC/* (1999)

95. Moser HG, Roussarie A. *Nucl. Instr. Meth. A* 384:491 (1997)

96. Barate R, et al (ALEPH Collaboration). *Eur. Phys. J. C* 7:553 (1999)

97. Abreu P, et al (DELPHI Collaboration). DELPHI 98-132 CONF 193 (1998)

98. Abbiendi G, et al (OPAL Collaboration). *Eur. Phys. J. C* 11:587 (1999)

99. Moore T, et al (SLD Collaboration). SLAC-R-551 (1999)

100. Abe K, et al (SLD Collaboration). SLAC-PUB-8225 (1999)

101. Buskulic D, et al (ALEPH Collaboration). *Phys. Lett.* B377:205 (1996)

102. Buskulic D, et al (ALEPH Collaboration). *Eur. Phys. J. C* 4:367 (1998)

103. Abreu P, et al (DELPHI Collaboration) DELPHI 98-131 CONF 192 (1998)

104. Abreu P, et al (DELPHI Collaboration) DELPHI 99-109 CONF 296.

105. Abe F, et al (CDF Collaboration). *Phys. Rev. Lett.* 82:3576 (1999)

106. Hagelin J.*Nucl. Phys.* B193:123 (1981); Voloshin M, et al *Sov. J. Nucl. Phys.* 46:112 (1987)

107. Abe F, et al (CDF Collaboration). *Phys. Rev. D* 59:032004 (1999)

108. Abreu P, et al (DELPHI Collaboration). DELPHI 99-109 CONF 296 (1999)

109. Acciarri M, et al (L3 Collaboration). *Phys. Lett.* B438:417 (1999)

110. Palla F, et al (ALEPH Collaboration). hep-ex/9905017 (1999)

111. Abe F, et al (CDF Collaboration). *Phys. Rev. Lett.* 77:1945 (1996)

112. Abe F, et al (CDF Collaboration). CDFNOTE 4672 (1998)